OM protein - protein search, using sw model

Run on: May 15, 2003, 13:23:56 ; Search time 35 Seconds
(without alignments)
167.515 Million cell updates/sec

Title: SEQID1_MOD
Perfect score: 197
Sequence: 1 ANSFLXXLRQGSLXRXCIXX.....XXAKXIFEDVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	179	90.9	44	20	AAY18300	Modified GLA domai
2	179	90.9	419	22	AAE08630	Human protein C de
3	179	90.9	419	22	AAB82677	Human protein C de
4	179	90.9	419	22	AAB82678	Human protein C de
5	176	89.3	44	20	AAY18301	Modified GLA domai
6	174	88.3	419	22	AAE08627	Human protein C de
7	174	88.3	419	22	AAE08628	Human protein C de
8	174	88.3	419	22	AAE08629	Human protein C de
9	174	88.3	419	22	AAB82675	Human protein C de
10	174	88.3	419	22	AAB82676	Human protein C de
11	173	87.8	44	20	AAY18298	Modified GLA domai
12	170	86.3	44	20	AAY18299	Modified GLA domai
13	168	85.3	44	20	AAY18307	Modified GLA domai
14	168	85.3	44	20	AAY18297	Modified GLA domai
15	160	81.2	44	20	AAY18309	Modified GLA domai
16	160	81.2	44	20	AAY18303	Human protein C GL
17	160	81.2	44	22	AAB36402	Human protein C ga
18	160	81.2	45	19	AAW75710	Partial human prot
19	160	81.2	415	21	AAY56803	Truncated human pr
20	160	81.2	419	14	AAR35760	Protein C (PC). H
21	160	81.2	419	19	AAW72753	Primary structure
22	160	81.2	419	22	AAE08625	Human mature wild
23	160	81.2	419	22	AAB82673	Wild-type human pr
24	160	81.2	419	22	AAB36894	Human protein C de
25	160	81.2	419	22	AAB36896	Human protein C de
26	160	81.2	419	22	AAB36897	Human protein C de
27	160	81.2	419	22	AAB36898	Human protein C de
28	160	81.2	419	23	AAU99002	Human Protein C zy
29	160	81.2	419	23	AAU99003	Human Protein C zy
30	160	81.2	419	23	AAU99004	Human Protein C zy
31	160	81.2	419	23	AAU99005	Human Protein C zy
32	160	81.2	419	23	AAU99006	Human Protein C zy
33	160	81.2	419	23	AAU99007	Human Protein C zy
34	160	81.2	419	23	AAU99008	Human Protein C zy
35	160	81.2	419	23	AAU99009	Human Protein C zy
36	160	81.2	419	23	AAU99010	Human Protein C zy
37	160	81.2	419	23	AAU99011	Human Protein C zy
38	160	81.2	419	23	AAU99012	Human Protein C zy
39	160	81.2	419	23	AAU99013	Human Protein C zy
40	160	81.2	419	23	AAU99014	Human Protein C zy
41	160	81.2	419	23	AAU99015	Human Protein C zy
42	160	81.2	419	23	AAU99016	Human Protein C zy
43	160	81.2	419	23	AAU99017	Human Protein C zy
44	160	81.2	419	23	AAU99018	Human Protein C zy
45	160	81.2	419	23	AAU99019	Human Protein C zy

ALIGNMENTS

RESULT 1
AAY18300
ID AAY18300 standard; peptide; 44 AA.
XX
AC AAY18300;
XX
DT 17-AUG-1999 (first entry)
XX
DE Modified GLA domain of vitamin K-dependent protein.
XX
KW GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..44
FT /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT acid"
XX
PN WO9920767-A1.
XX
PD 29-APR-1999.
XX
PF 20-OCT-1998; 98WO-US22152.
XX
PR 23-OCT-1997; 97US-0955636.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Nelsestuen GL;
XX

```
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT   acid domain, useful for treating clotting disorders
XX
PS   Claim 9; Page 79; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein.
XX
SQ   Sequence   44 AA;

  Query Match             90.9%;  Score 179;  DB 20;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-22;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||||||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 2
AAE08630
ID   AAE08630 standard; Protein; 419 AA.
XX
AC   AAE08630;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #4.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US01221.
XX
PR   11-FEB-2000; 2000US-0181948.
PR   14-MAR-2000; 2000US-0189199.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15228.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions
XX
PS   Claim 6; Page 50-51; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and
CC   increased sensitivity to thrombin activation compared to wild type
CC   protein C, and retains the biological activity of the wild type human
CC   protein C. Protein C derivatives are useful in the manufacture of a
CC   medicament for the treatment of acute coronary syndromes e.g. myocardial
CC   infarction and unstable angina; and disease states predisposing to
CC   thrombosis; vascular occlusive disorders and hypercoagulable states e.g.
CC   disseminated intravascular coagulation (DIC), burns, transplantations,
CC   thalassaemia, sickle cell disease, viral haemorrhagic fever and
CC   haemolytic uremic syndrome; sepsis in combination with bacterial
CC   permeability increasing protein; thrombotic disorders in combination
CC   with an anti-platelet agent; protein C deficiency; acute arterial
CC   thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral
CC   or peripheral arteries or in vascular grafts in combination with a
CC   thrombolytic agent. Nucleic acid molecules of the invention are useful
CC   for treating humans with genetically predisposed prothrombotic disorders
CC   by gene therapy. The present sequence is human protein C derivative.
XX
SQ   Sequence   419 AA;

  Query Match             90.9%;  Score 179;  DB 22;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 1.4e-21;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  ||||||  | ||  ||||  || ||||||||||||||
Db      1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 3
AAB82677
ID   AAB82677 standard; Protein; 419 AA.
XX
AC   AAB82677;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (H10Q/S11G/Q32E/N33D/L194S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder;
KW   hypercoagulation; sepsis; protein C deficiency; occlusion;
KW   thromboembolism; stenosis; antibacterial; immunosuppressive;
KW   thrombolytic; cardiant; antianginal; anticoagulant; therapy;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "His in wild-type protein"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Domain          1..45
FT                   /note= "Gla domain"
FT   Disulfide-bond  50..69
FT   Disulfide-bond  59..64
FT   Disulfide-bond  80..89
FT   Disulfide-bond  98..109
FT   Disulfide-bond  120..133
FT   Disulfide-bond  141..277
FT   Disulfide-bond  196..212
FT   Disulfide-bond  331..345
FT   Disulfide-bond  356..384
FT   Cleavage-site   156..157
FT                   /note= "cleavage makes a 2-chain inactive
FT                           precursor (155-amino acid light chain
FT                           attached via a disulfide bond to a
```

```
FT                              262-amino acid heavy chain)"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Peptide         158..169
FT                   /note= "activation peptide; removal activates the
FT                          2-chain zymogen"
FT   Cleavage-site   169..170
FT                   /note= "thrombin cleavage site"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US00020.
XX
PR   02-FEB-2000; 2000US-0179801.
PR   14-MAR-2000; 2000US-0189197.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26365.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute
PT   arterial thrombotic occlusion, and thromboembolism -
XX
PS   Claim 5; Page 54-55; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative
CC   in which His at position 10 of the mature wild-type protein C
CC   sequence (see AAB82673) is substituted with Gln, Ser at position 11
CC   with Gly, Gln at position 32 with Glu, Asn at position 33 with Asp,
CC   and Leu at position 194 with Ser.  It is an example of protein C
CC   derivatives of the invention that have at least 2 amino acid
CC   substitutions, but which have increased anticoagulant activity and
CC   resistance to inactivation by serpins compared with the wild-type
CC   protein, while retaining the biological activity of the wild-type
CC   protein.  A method of producing the derivatives using recombinant
CC   DNA methods is claimed.  The protein C derivatives are useful for
CC   treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina),
CC   vascular occlusive disorders and hypercoagulable states, sepsis (in
CC   combination with bactericidal permeability increasing protein or
CC   with tissue factor pathway inhibitor), thrombotic disorders (in
CC   combination with an anti-platelet agent or by local delivery through
CC   an intracoronary catheter), protein C deficiency, acute arterial
CC   thrombotic occlusion, thromboembolism, or stenosis in coronary,
CC   cerebral or peripheral arteries or in vascular grafts.  Human
CC   patients with genetically predisposed prothrombotic disorders may
CC   be treated by gene therapy (all claimed).
XX
SQ   Sequence   419 AA;

  Query Match             90.9%;  Score 179;  DB 22;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 1.4e-21;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  |||||| | ||  ||||  || ||||||||||||||||
Db       1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 4
AAB82678
ID   AAB82678 standard; Protein; 419 AA.
XX
AC   AAB82678;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (H10Q/S11G/Q32E/N33D/L194S/T254S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder;
KW   hypercoagulation; sepsis; protein C deficiency; occlusion;
KW   thromboembolism; stenosis; antibacterial; immunosuppressive;
KW   thrombolytic; cardiant; antianginal; anticoagulant; therapy;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "His in wild-type protein"
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Misc-difference 254
FT                   /note= "Thr in wild-type protein"
FT   Domain          1..45
FT                   /note= "Gla domain"
FT   Disulfide-bond  50..69
FT   Disulfide-bond  59..64
FT   Disulfide-bond  80..89
FT   Disulfide-bond  98..109
FT   Disulfide-bond  120..133
FT   Disulfide-bond  141..277
FT   Disulfide-bond  196..212
FT   Disulfide-bond  331..345
FT   Disulfide-bond  356..384
FT   Cleavage-site   156..157
FT                   /note= "cleavage makes a 2-chain inactive
FT                          precursor (155-amino acid light chain
FT                          attached via a disulfide bond to a
FT                          262-amino acid heavy chain)"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
```

```
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Peptide         158..169
FT                   /note= "activation peptide; removal activates the
FT                         2-chain zymogen"
FT   Cleavage-site   169..170
FT                   /note= "thrombin cleavage site"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US00020.
XX
PR   02-FEB-2000; 2000US-0179801.
PR   14-MAR-2000; 2000US-0189197.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute
PT   arterial thrombotic occlusion, and thromboembolism -
XX
PS   Claim 6; Page 56-57; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative
CC   in which His at position 10 of the wild-type protein C sequence (see
CC   AAB82673) is substituted with Gln, Ser at position 11 with Gly, Gln
CC   at position 32 with Glu, Asn at position 33 with Asp, Leu at position
CC   194 with Ser, and Thr at position 254 with Ser.  It is an example of
CC   protein C derivatives of the invention that have at least 2 amino acid
CC   substitutions, but which have increased anticoagulant activity and
CC   resistance to inactivation by serpins compared with the wild-type
CC   protein, while retaining the biological activity of the wild-type
CC   protein.  A method of producing the derivatives using recombinant
CC   DNA methods is claimed.  The protein C derivatives are useful for
CC   treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina),
CC   vascular occlusive disorders and hypercoagulable states, sepsis (in
CC   combination with bactericidal permeability increasing protein or
CC   with tissue factor pathway inhibitor), thrombotic disorders (in
CC   combination with an anti-platelet agent or by local delivery through
CC   an intracoronary catheter), protein C deficiency, acute arterial
CC   thrombotic occlusion, thromboembolism, or stenosis in coronary,
CC   cerebral or peripheral arteries or in vascular grafts.  Human
CC   patients with genetically predisposed prothrombotic disorders may
CC   be treated by gene therapy (all claimed).
XX
SQ   Sequence   419 AA;

  Query Match             90.9%;  Score 179;  DB 22;  Length 419;
  Best Local Similarity   79.5%;  Pred. No. 1.4e-21;
  Matches   35;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  |||||| | ||  |||| || ||||||||||||||||
Db      1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 5
AAY18301
ID   AAY18301 standard; peptide; 44 AA.
XX
AC   AAY18301;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                         acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US22152.
XX
PR   23-OCT-1997;   97US-0955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT   acid domain, useful for treating clotting disorders
XX
PS   Claim 9; Page 82; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein.
XX
SQ   Sequence   44 AA;

  Query Match             89.3%;  Score 176;  DB 20;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 4.1e-22;
  Matches   43;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||:|||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLREGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 6
AAE08627
ID   AAE08627 standard; Protein; 419 AA.
XX
AC   AAE08627;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #1.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
```

```
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US01221.
XX
PR   11-FEB-2000; 2000US-0181948.
PR   14-MAR-2000; 2000US-0189199.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15225.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions  -
XX
PS   Claim 3; Page 46-47; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and
CC   increased sensitivity to thrombin activation compared to wild type
CC   protein C, and retains the biological activity of the wild type human
CC   protein C. Protein C derivatives are useful in the manufacture of a
CC   medicament for the treatment of acute coronary syndromes e.g. myocardial
CC   infarction and unstable angina; and disease states predisposing to
CC   thrombosis; vascular occlusive disorders and hypercoagulable states e.g.
CC   disseminated intravascular coagulation (DIC), burns, transplantations,
CC   thalassaemia, sickle cell disease, viral haemorrhagic fever and
CC   haemolytic uremic syndrome; sepsis in combination with bacterial
CC   permeability increasing protein; thrombotic disorders in combination
CC   with an anti-platelet agent; protein C deficiency; acute arterial
CC   thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral
CC   or peripheral arteries or in vascular grafts in combination with a
CC   thrombolytic agent. Nucleic acid molecules of the invention are useful
CC   for treating humans with genetically predisposed prothrombotic disorders
CC   by gene therapy. The present sequence is human protein C derivative.
XX
SQ   Sequence   419 AA;

  Query Match             88.3%;  Score 174;  DB 22;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  || ||| | ||  ||||  || ||||||||||||||||
Db       1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 7
AAE08628
ID   AAE08628 standard; Protein; 419 AA.
XX
AC   AAE08628;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #2.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
```

```
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US01221.
XX
PR   11-FEB-2000; 2000US-0181948.
PR   14-MAR-2000; 2000US-0189199.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15226.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions  -
XX
PS   Claim 4; Page 47-48; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and
CC   increased sensitivity to thrombin activation compared to wild type
CC   protein C, and retains the biological activity of the wild type human
CC   protein C. Protein C derivatives are useful in the manufacture of a
CC   medicament for the treatment of acute coronary syndromes e.g. myocardial
CC   infarction and unstable angina; and disease states predisposing to
CC   thrombosis; vascular occlusive disorders and hypercoagulable states e.g.
CC   disseminated intravascular coagulation (DIC), burns, transplantations,
CC   thalassaemia, sickle cell disease, viral haemorrhagic fever and
CC   haemolytic uremic syndrome; sepsis in combination with bacterial
CC   permeability increasing protein; thrombotic disorders in combination
CC   with an anti-platelet agent; protein C deficiency; acute arterial
CC   thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral
CC   or peripheral arteries or in vascular grafts in combination with a
CC   thrombolytic agent. Nucleic acid molecules of the invention are useful
CC   for treating humans with genetically predisposed prothrombotic disorders
CC   by gene therapy. The present sequence is human protein C derivative.
XX
SQ   Sequence   419 AA;

  Query Match             88.3%;  Score 174;  DB 22;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  || ||| | ||  ||||  || ||||||||||||||||
Db       1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 8
AAE08629
ID   AAE08629 standard; Protein; 419 AA.
XX
AC   AAE08629;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Human protein C derivative #3.
XX
KW   Human; protein C derivative; anticoagulation activity; thrombosis;
```

```
KW   serpin inactivation; acute coronary syndrome; myocardial infarction;
KW   vascular occlusive disorder; hypercoagulable state; angina; sepsis;
KW   disseminated intravascular coagulation; DIC; burn; transplantation;
KW   sickle cell disease; viral haemorrhagic fever; protein C deficiency;
KW   haemolytic uremic syndrome; acute arterial thrombotic occlusion;
KW   thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "Encoded by CAA"
XX
PN   WO200159084-A1.
XX
PD   16-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US01221.
XX
PR   11-FEB-2000; 2000US-0181948.
PR   14-MAR-2000; 2000US-0189199.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Grinnell BW,  Jones BE;
XX
DR   WPI; 2001-514662/56.
DR   N-PSDB; AAD15227.
XX
PT   Protein C derivative for treating acute coronary syndromes, vascular
PT   occlusive disorders, thrombotic disorders and sepsis, comprises
PT   substitutions at specified amino acid positions  -
XX
PS   Claim 5; Page 48-49; 59pp; English.
XX
CC   The invention relates to human protein C derivatives and nucleic acid
CC   molecules encoding such derivatives. These derivatives have increased
CC   anticoagulation activity, resistance to serpin inactivation and
CC   increased sensitivity to thrombin activation compared to wild type
CC   protein C, and retains the biological activity of the wild type human
CC   protein C. Protein C derivatives are useful in the manufacture of a
CC   medicament for the treatment of acute coronary syndromes e.g. myocardial
CC   infarction and unstable angina; and disease states predisposing to
CC   thrombosis; vascular occlusive disorders and hypercoagulable states e.g.
CC   disseminated intravascular coagulation (DIC), burns, transplantations,
CC   thalassaemia, sickle cell disease, viral haemorrhagic fever and
CC   haemolytic uremic syndrome; sepsis in combination with bacterial
CC   permeability increasing protein; thrombotic disorders in combination
CC   with an anti-platelet agent; protein C deficiency; acute arterial
CC   thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral
CC   or peripheral arteries or in vascular grafts in combination with a
CC   thrombolytic agent. Nucleic acid molecules of the invention are useful
CC   for treating humans with genetically predisposed prothrombotic disorders
CC   by gene therapy. The present sequence is human protein C derivative.
XX
SQ   Sequence   419 AA;

  Query Match             88.3%;  Score 174;  DB 22;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  || ||| | ||  ||||  || ||||||||||||||||
Db       1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 9
AAB82675
ID   AAB82675 standard; Protein; 419 AA.
XX
AC   AAB82675;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (S11G/Q32E/N33D/L194S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder;
KW   hypercoagulation; sepsis; protein C deficiency; occlusion;
KW   thromboembolism; stenosis; antibacterial; immunosuppressive;
KW   thrombolytic; cardiant; antianginal; anticoagulant; therapy;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Domain          1..45
FT                   /note= "Gla domain"
FT   Disulfide-bond  50..69
FT   Disulfide-bond  59..64
FT   Disulfide-bond  80..89
FT   Disulfide-bond  98..109
FT   Disulfide-bond  120..133
FT   Disulfide-bond  141..277
FT   Disulfide-bond  196..212
FT   Disulfide-bond  331..345
FT   Disulfide-bond  356..384
FT   Cleavage-site   156..157
FT                   /note= "cleavage makes a 2-chain inactive
FT                          precursor (155-amino acid light chain
FT                          attached via a disulfide bond to a
FT                          262-amino acid heavy chain)"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Peptide         158..169
FT                   /note= "activation peptide; removal activates the
FT                          2-chain zymogen"
FT   Cleavage-site   169..170
FT                   /note= "thrombin cleavage site"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US00020.
```

```
XX
PR   02-FEB-2000; 2000US-0179801.
PR   14-MAR-2000; 2000US-0189197.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26363.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute
PT   arterial thrombotic occlusion, and thromboembolism -
XX
PS   Claim 3; Page 52-53; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C
CC   derivative in which Ser at amino acid position 11 of the mature
CC   wild-type protein C sequence (see AAB82673) is substituted with
CC   Gly, Gln at position 32 with Glu, Asn at position 33 with Asp, and
CC   Leu at position 194 with Ser.  The protein is an example of protein
CC   C derivatives of the invention that have at least 2 amino acid
CC   substitutions, but which have increased anticoagulant activity and
CC   resistance to inactivation by serpins compared with the wild-type
CC   protein, while retaining the biological activity of the wild-type
CC   protein.  A method of producing the derivatives using recombinant
CC   DNA methods is claimed.  The protein C derivatives are useful for
CC   treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina),
CC   vascular occlusive disorders and hypercoagulable states, sepsis (in
CC   combination with bactericidal permeability increasing protein or
CC   with tissue factor pathway inhibitor), thrombotic disorders (in
CC   combination with an anti-platelet agent or by local delivery through
CC   an intracoronary catheter), protein C deficiency, acute arterial
CC   thrombotic occlusion, thromboembolism, or stenosis in coronary,
CC   cerebral or peripheral arteries or in vascular grafts.  Human
CC   patients with genetically predisposed prothrombotic disorders may
CC   be treated by gene therapy (all claimed).
XX
SQ   Sequence   419 AA;

  Query Match             88.3%;  Score 174;  DB 22;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  || ||| | ||  ||||  || ||||||||||||||||
Db       1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 10
AAB82676
ID   AAB82676 standard; Protein; 419 AA.
XX
AC   AAB82676;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Human protein C derivative (S11G/Q32E/N33D/L194S/T254S).
XX
KW   Protein C; human; coronary syndrome; thrombosis; angina;
KW   myocardial infarction; vascular occlusive disorder;
KW   hypercoagulation; sepsis; protein C deficiency; occlusion;
KW   thromboembolism; stenosis; antibacterial; immunosuppressive;
KW   thrombolytic; cardiant; antianginal; anticoagulant; therapy;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 11
FT                   /note= "Ser in wild-type protein"
FT   Misc-difference 32
FT                   /note= "Gln in wild-type protein"
FT   Misc-difference 33
FT                   /note= "Asn in wild-type protein"
FT   Misc-difference 194
FT                   /note= "Leu in wild-type protein"
FT   Misc-difference 254
FT                   /note= "Thr in wild-type protein"
FT   Domain          1..45
FT                   /note= "Gla domain"
FT   Disulfide-bond  50..69
FT   Disulfide-bond  59..64
FT   Disulfide-bond  80..89
FT   Disulfide-bond  98..109
FT   Disulfide-bond  120..133
FT   Disulfide-bond  141..277
FT   Disulfide-bond  196..212
FT   Disulfide-bond  331..345
FT   Disulfide-bond  356..384
FT   Cleavage-site   156..157
FT                   /note= "cleavage makes a 2-chain inactive
FT                          precursor (155-amino acid light chain
FT                          attached via a disulfide bond to a
FT                          262-amino acid heavy chain)"
FT   Modified-site   6
FT                   /note= "gamma-carboxylated"
FT   Modified-site   7
FT                   /note= "gamma-carboxylated"
FT   Modified-site   14
FT                   /note= "gamma-carboxylated"
FT   Modified-site   16
FT                   /note= "gamma-carboxylated"
FT   Modified-site   19
FT                   /note= "gamma-carboxylated"
FT   Modified-site   20
FT                   /note= "gamma-carboxylated"
FT   Modified-site   25
FT                   /note= "gamma-carboxylated"
FT   Modified-site   26
FT                   /note= "gamma-carboxylated"
FT   Peptide         158..169
FT                   /note= "activation peptide; removal activates the
FT                          2-chain zymogen"
FT   Cleavage-site   169..170
FT                   /note= "thrombin cleavage site"
FT   Modified-site   29
FT                   /note= "N-glycosylated"
FT   Modified-site   248
FT                   /note= "N-glycosylated"
FT   Modified-site   313
FT                   /note= "N-glycosylated"
FT   Modified-site   329
FT                   /note= "N-glycosylated"
XX
PN   WO200157193-A2.
XX
PD   09-AUG-2001.
XX
PF   19-JAN-2001; 2001WO-US00020.
XX
PR   02-FEB-2000; 2000US-0179801.
PR   14-MAR-2000; 2000US-0189197.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Gerlitz BE,  Jones BE;
XX
DR   WPI; 2001-496919/54.
DR   N-PSDB; AAH26364.
XX
PT   Novel human protein C derivative for treating, e.g., myocardial
PT   infarction, unstable angina, sepsis, thrombotic disorders, acute
```

```
PT   arterial thrombotic occlusion, and thromboembolism -
XX
PS   Claim 4; Page 53-54; 63pp; English.
XX
CC   The present sequence is that of a claimed human protein C derivative
CC   in which Ser at position 11 of the mature wild-type protein C
CC   sequence (see AAB82673) is substituted with Gly, Gln at position 32
CC   with Glu, Asn at position 33 with Asp, Leu at position 194 with Ser,
CC   and Thr at position 254 with Ser.  It is an example of protein C
CC   derivatives of the invention that have at least 2 amino acid
CC   substitutions, but which have increased anticoagulant activity and
CC   resistance to inactivation by serpins compared with the wild-type
CC   protein, while retaining the biological activity of the wild-type
CC   protein.  A method of producing the derivatives using recombinant
CC   DNA methods is claimed.  The protein C derivatives are useful for
CC   treating coronary syndromes and disease states predisposing to
CC   thrombosis (e.g. myocardial infarction and unstable angina),
CC   vascular occlusive disorders and hypercoagulable states, sepsis (in
CC   combination with bactericidal permeability increasing protein or
CC   with tissue factor pathway inhibitor), thrombotic disorders (in
CC   combination with an anti-platelet agent or by local delivery through
CC   an intracoronary catheter), protein C deficiency, acute arterial
CC   thrombotic occlusion, thromboembolism, or stenosis in coronary,
CC   cerebral or peripheral arteries or in vascular grafts.  Human
CC   patients with genetically predisposed prothrombotic disorders may
CC   be treated by gene therapy (all claimed).
XX
SQ   Sequence  419 AA;

  Query Match             88.3%;  Score 174;  DB 22;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 1e-20;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  || ||| | ||  ||||  || ||||||||||||||||
Db      1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 11
AAY18298
ID   AAY18298 standard; peptide; 44 AA.
XX
AC   AAY18298;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                          acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US22152.
XX
PR   23-OCT-1997;   97US-0955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT   acid domain, useful for treating clotting disorders
XX
PS   Claim 7; Page 78; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein.
XX
SQ   Sequence  44 AA;

  Query Match             87.8%;  Score 173;  DB 20;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 1.3e-21;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||| ||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRQSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 12
AAY18299
ID   AAY18299 standard; peptide; 44 AA.
XX
AC   AAY18299;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                          acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US22152.
XX
PR   23-OCT-1997;   97US-0955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT   acid domain, useful for treating clotting disorders
XX
PS   Claim 8; Page 78; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein.
```

```
XX
SQ   Sequence   44 AA;

  Query Match             86.3%;  Score 170;  DB 20;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 4.2e-21;
  Matches   42;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||: |||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRESSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 13
AAY18307
ID   AAY18307 standard; peptide; 44 AA.
XX
AC   AAY18307;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                          acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US22152.
XX
PR   23-OCT-1997;   97US-0955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT   acid domain, useful for treating clotting disorders
XX
PS   Disclosure; Page 79; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein.
XX
SQ   Sequence   44 AA;

  Query Match             85.3%;  Score 168;  DB 20;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 9.1e-21;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  |||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRDSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 14
AAY18297
ID   AAY18297 standard; peptide; 44 AA.
XX
AC   AAY18297;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                          acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US22152.
XX
PR   23-OCT-1997;   97US-0955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT   acid domain, useful for treating clotting disorders
XX
PS   Claim 6; Page 78; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein.
XX
SQ   Sequence   44 AA;

  Query Match             85.3%;  Score 168;  DB 20;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 9.1e-21;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  |||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 15
AAY18309
ID   AAY18309 standard; peptide; 44 AA.
XX
AC   AAY18309;
XX
DT   17-AUG-1999  (first entry)
XX
DE   Modified GLA domain of vitamin K-dependent protein.
XX
KW   GLA domain; mutein; vitamin K-dependent protein; clotting disorder;
KW   therapy.
XX
OS   Homo sapiens.
```

```
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..44
FT                   /note= "Xaa= gamma-carboxyglutamic acid, or glutamic
FT                          acid"
XX
PN   WO9920767-A1.
XX
PD   29-APR-1999.
XX
PF   20-OCT-1998;   98WO-US22152.
XX
PR   23-OCT-1997;   97US-0955636.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Nelsestuen GL;
XX
DR   WPI; 1999-288309/24.
XX
PT   Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT   acid domain, useful for treating clotting disorders
XX
PS   Disclosure; Page 79-80; 86pp; English.
XX
CC   This sequence represents a modified GLA (gamma-carboxyglutamic acid)
CC   domain. The invention relates to a vitamin K-dependent polypeptide
CC   comprising a modified GLA domain containing an amino acid substitution
CC   which enhances membrane binding of the modified polypeptide as compared
CC   to the native polypeptide. The polypeptide is used to treat a clotting
CC   disorder by decreasing or increasing clot formation. Modification of the
CC   GLA domain results in a protein which has enhanced membrane binding
CC   affinity as compared to the native protein.
XX
SQ   Sequence   44 AA;

  Query Match             81.2%;  Score 160;  DB 20;  Length 44;
  Best Local Similarity   93.2%;  Pred. No. 2e-19;
  Matches   41;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  ||||||||||||||| |||||||||||||||
Db      1 ANSFLXXLRHSSLXRXCIXXICDFXXAFXIFEDVDDTLAFWSKH 44


Search completed: May 15, 2003, 13:27:16
Job time : 36 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 15, 2003, 13:26:07 ; Search time 18 Seconds
(without alignments)
234.995 Million cell updates/sec

Title: SEQID1_MOD
Perfect score: 197
Sequence: 1 ANSFLXXLRQGSLXRXCIXX.....XXAKXIFEDVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	160	81.2	461	1	KXHU	protein C (activat
2	140	71.1	461	1	JX0210	protein C (activat
3	139	70.6	461	1	S18994	protein C (activat
4	122	61.9	456	1	KXBO	protein C (activat
5	115	58.4	482	1	EXRT	coagulation factor
6	114	57.9	492	1	EXBO	coagulation factor
7	110	55.8	488	1	EXHU	coagulation factor
8	101	51.3	443	2	I46932	coagulation factor
9	99	50.3	466	1	KFHU7	coagulation factor
10	86.5	43.9	617	2	S10511	thrombin (EC 3.4.2
11	86.5	43.9	618	2	A35827	thrombin (EC 3.4.2
12	86	43.7	475	1	EXCH	coagulation factor
13	85	43.1	407	1	KFBO7	coagulation factor
14	85	43.1	642	2	S53434	plasma protein S p
15	85	43.1	676	1	KXHUS	plasma protein S p
16	84	42.6	622	1	TBHU	thrombin (EC 3.4.2
17	81	41.1	646	2	S38819	plasma protein S -
18	80	40.6	452	1	A30351	coagulation factor
19	80	40.6	459	2	JQ0419	coagulation factor
20	80	40.6	461	1	KFHU	coagulation factor
21	80	40.6	675	1	KXBOS	plasma protein S p
22	78	39.6	642	2	S53433	plasma protein S p
23	78	39.6	675	1	KXRTS	plasma protein S p
24	73	37.1	416	1	KFBO	coagulation factor
25	72	36.5	625	1	TBBO	thrombin (EC 3.4.2
26	71	36.0	675	1	KXMSS	plasma protein S p
27	69.5	35.3	396	1	KXBOZ	plasma protein Z -
28	65.5	33.2	422	1	KXHUZ	plasma protein Z p
29	65	33.0	673	2	A48089	growth arrest-spec
30	64	32.5	674	2	I55476	growth potentiatin
31	63	32.0	678	2	B48089	growth arrest-spec
32	56.5	28.7	594	2	D84859	probable MAP kinas
33	54.5	27.7	603	2	C96575	probable MAP kinas
34	53.5	27.2	576	2	G96763	probable MAP kinas
35	53	26.9	606	2	T40556	hypothetical prote
36	50	25.4	1684	2	T02367	hypothetical prote
37	49	24.9	1363	2	I58375	protein-tyrosine k
38	48.5	24.6	323	2	T25948	hypothetical prote
39	48.5	24.6	510	2	E82918	ammonium transport
40	48	24.4	422	2	T39306	mitogen-activated
41	48	24.4	1235	2	D32433	VSG expression sit
42	48	24.4	1298	2	A48999	protein-tyrosine k
43	47.5	24.1	1089	1	S33727	platelet-derived g
44	47	23.9	182	2	JC1189	tyrosine kinase re
45	47	23.9	245	1	NDECR5	type II site-speci

ALIGNMENTS

RESULT 1
KXHU
protein C (activated) (EC 3.4.21.69) precursor - human
N;Alternate names: autoprothrombin IIA; plasma protein C
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C;Accession: A22331; A25426; A21781; A23789; A00927
R;Foster, D.C.; Yoshitake, S.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 4673-4677, 1985
A;Title: The nucleotide sequence of the gene for human protein C.
A;Reference number: A22331; MUID:85270390; PMID:2991887
A;Accession: A22331
A;Molecule type: DNA
A;Residues: 1-461 <FOS1>
A;Cross-references: GB:M11228; NID:g190333; PIDN:AAA60166.1; PID:g190334
R;Plutzky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.
Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986
A;Title: Evolution and organization of the human protein C gene.
A;Reference number: A25426; MUID:86120978; PMID:3511471
A;Accession: A25426
A;Molecule type: DNA
A;Residues: 1-445,'L',446-461 <PLU>
A;Cross-references: GB:M12712; NID:g190330; PIDN:AAA60165.1; PID:g190332
R;Foster, D.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 4766-4770, 1984
A;Title: Characterization of a cDNA coding for human protein C.
A;Reference number: A21781; MUID:84272714; PMID:6589623
A;Accession: A21781
A;Molecule type: mRNA
A;Residues: 'Q',107-461 <FOS2>
A;Cross-references: GB:K02059; NID:g190322; PIDN:AAA60164.1; PID:g190323
R;Beckmann, R.J.; Schmidt, R.J.; Santerre, R.F.; Plutzky, J.; Crabtree, G.R.; Long, G
Nucleic Acids Res. 13, 5233-5247, 1985
A;Title: The structure and evolution of a 461 amino acid human protein C precursor an
A;Reference number: A23789; MUID:85269639; PMID:2991859
A;Accession: A23789
A;Molecule type: mRNA
A;Residues: 1-461 <BEC>
A;Cross-references: GB:X02750; NID:g35689; PIDN:CAA26528.1; PID:g763120
R;Miletich, J.P.; Broze Jr., G.J.
J. Biol. Chem. 265, 11397-11404, 1990
A;Title: Beta protein C is not glycosylated at asparagine 329. The rate of translatio
A;Reference number: A44605; MUID:90293094; PMID:1694179
A;Contents: annotation; carbohydrate binding sites; activation peptide
A;Note: the alpha form of protein C is glycosylated at Asn-329, and the beta form is
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A;Title: O-linked fucose is present in the first epidermal growth factor domain of fa
A;Reference number: A44606; MUID:92184750; PMID:1544894
A;Contents: annotation; beta-hydroxyaspartic acid
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that
ivation of factor Va is strongly enhanced by complexing with protein S. Protein C als

```
C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c
bin, which cleaves a dodecapeptide from the amino end of the heavy chain; this reaction,
C;Genetics:
A;Gene: GDB:PROC
A;Cross-references: GDB:120317; OMIM:176860
A;Map position: 2q13-2q21
A;Introns: 24/1; 79/3; 88/1; 134/1; 179/1; 226/3; 266/1
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
F;1-32/Domain: signal sequence #status predicted <SIG>
F;27-86/Domain: Gla domain homology <GLA>
F;33-42/Domain: propeptide #status predicted <PRO>
F;43-197/Product: protein C light chain #status predicted <LCH>
F;92-131/Domain: EGF homology <EG1>
F;140-175/Domain: EGF homology <EG2>
F;200-461/Product: protein C heavy chain #status predicted <HCH>
F;200-211/Domain: activation peptide #status experimental <APT>
F;212-445/Domain: trypsin homology <TRY>
F;48,49,56,58,61,62,67,68,71/Modified site: gamma-carboxyglutamic acid (Glu) #status exp
F;59-64,92-105,101-120,122-131,140-151,147-160,162-175,183-319,238-254,373-387,398-426/D
F;106-111/Disulfide bonds: #status predicted
F;110/Binding site: carbohydrate (Thr) (covalent) #status absent
F;113/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;139,290,355/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;211-212/Cleavage site: Arg-Leu (thrombin) #status experimental
F;253,299,402/Active site: His, Asp, Ser #status predicted
F;371/Binding site: carbohydrate (Asn) (covalent) (partial) #status atypical

  Query Match             81.2%;  Score 160;  DB 1;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 8.8e-18;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  ||  || | ||  ||||  || ||::|||||||||||
Db     43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 2
JX0210
protein C (activated) (EC 3.4.21.69) precursor - mouse
N;Alternate names: vitamin K-dependent serine proteinase
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: JX0210
R;Tada, N.; Sato, M.; Tsujimura, A.; Iwase, R.; Hashimoto-Gotoh, T.
J. Biochem. 111, 491-495, 1992
A;Title: Isolation and characterization of a mouse protein C cDNA.
A;Reference number: JX0210; MUID:92316897; PMID:1618739
A;Accession: JX0210
A;Molecule type: mRNA
A;Residues: 1-461 <TAD>
A;Cross-references: GB:D10445; NID:g220385; PIDN:BAA01235.1; PID:g220386
A;Experimental source: liver
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re
s.
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-33/Domain: signal sequence #status predicted <SIG>
F;27-85/Domain: Gla domain homology <GLA>
F;34-41/Domain: propeptide #status predicted <PRO>
F;42-196,199-461/Product: protein C #status predicted <PRC>
F;42-196/Domain: light chain #status predicted <PCL>
F;91-130/Domain: EGF homology <EG1>
F;139-174/Domain: EGF homology <EG2>
F;199-461/Domain: heavy chain #status predicted <PCH>
F;199-211/Domain: activation peptide #status predicted <ACT>
F;212-461/Product: vitamin K-dependent serine proteinase #status predicted <VIT>
F;212-445/Domain: trypsin homology <TRY>
F;47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;121-130,139-150,146-159,161-174,182-319,238-254,373-387,398-426/Disulfide bonds: #stat
F;214,290,355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;253,299,402/Active site: His, Asp, Ser #status predicted

  Query Match             71.1%;  Score 140;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 1.5e-14;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  :| ||| | |:  ||||  |: ||::|:|||||| |:
Db     42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
S18994
protein C (activated) (EC 3.4.21.69) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Oct-1999
C;Accession: S18994; S24312
R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.
submitted to the EMBL Data Library, February 1992
A;Description: The cDNA cloninig and mRNA expression of rat protein C.
A;Reference number: S18994
A;Accession: S18994
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <OKA>
A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963
R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.
Biochim. Biophys. Acta 1131, 329-332, 1992
A;Title: The cDNA cloning and mRNA expression of rat protein C.
A;Reference number: S24312; MUID:92329550; PMID:1627650
A;Accession: S24312
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <OKA2>
A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
C;Keywords: beta-hydroxyaspartic acid; glycoprotein; hydrolase; serine proteinase
F;1-32/Domain: signal sequence #status predicted <SIG>
F;27-85/Domain: Gla domain homology <GLA>
F;33-42/Domain: propeptide #status predicted <PRO>
F;43-461/Product: protein C #status predicted <PRC>
F;91-130/Domain: EGF homology <EG1>
F;139-174/Domain: EGF homology <EG2>
F;213-445/Domain: trypsin homology <TRY>
F;47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;121-130,139-150,146-159,161-174,182-320,239-255,373-387,398-426/Disulfide bonds: #s
F;215,291,355/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;254,300,402/Active site: His, Asp, Ser #status predicted

  Query Match             70.6%;  Score 139;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 2.1e-14;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  :| ||| | |:  ||||  |: ||::|:|||||| |:
Db     42 ANSFLEEVRAGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
KXBO
protein C (activated) (EC 3.4.21.69) precursor - bovine (fragment)
N;Alternate names: autoprothrombin IIA; plasma protein C
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Nov-1980 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C;Accession: A26250; A18385; A18386; A00928
R;Long, G.L.; Balagaje, R.M.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 81, 5653-5656, 1984
A;Title: Cloning and sequence of liver cDNA coding for bovine protein C.
A;Reference number: A26250; MUID:85014826; PMID:6091100
A;Accession: A26250
A;Molecule type: mRNA
A;Residues: 1-456 <LON>
R;Fernlund, P.; Stenflo, J.
```

```
J. Biol. Chem. 257, 12170-12179, 1982
A;Title: Amino acid sequence of the light chain of bovine protein C.
A;Reference number: A18385; MUID:83007325; PMID:6896876
A;Accession: A18385
A;Molecule type: protein
A;Residues: 40-194 <FER>
A;Note: 82-Lys was also found
R;Drakenberg, T.; Fernlund, P.; Roepstorff, P.; Stenflo, J.
Proc. Natl. Acad. Sci. U.S.A. 80, 1802-1806, 1983
A;Title: beta-Hydroxyaspartic acid in vitamin K-dependent protein C.
A;Reference number: A19316; MUID:83169769; PMID:6572939
A;Contents: annotation; revision to residue 110
R;Stenflo, J.; Fernlund, P.
J. Biol. Chem. 257, 12180-12190, 1982
A;Title: Amino acid sequence of the heavy chain of bovine protein C.
A;Reference number: A18386; MUID:83007326; PMID:6896877
A;Accession: A18386
A;Molecule type: protein
A;Residues: 197-454,'PV' <STE>
R;Esmon, N.L.; DeBault, L.E.; Esmon, C.T.
J. Biol. Chem. 258, 5548-5553, 1983
A;Title: Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless p
A;Reference number: A37541; MUID:83213513; PMID:6304092
A;Contents: annotation; activation; calcium binding
R;Johnson, A.E.; Esmon, N.L.; Laue, T.M.; Esmon, C.T.
J. Biol. Chem. 258, 5554-5560, 1983
A;Title: Structural changes required for activation of protein C are induced by Ca2+ bin
A;Reference number: A37542; MUID:83213514; PMID:6406503
A;Contents: annotation; activation; calcium binding
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re
s.
C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c
bin, which cleaves a tetradecapeptide from the amino end of the heavy chain; this reacti
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with stro
cognition of the thrombin-thrombomodulin complex.
C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
F;1-29/Domain: signal sequence (fragment) #status predicted <SIG>
F;24-83/Domain: Gla domain homology <GLA>
F;30-39/Domain: propeptide #status predicted <PRO>
F;40-194/Product: protein C light chain #status experimental <LCH>
F;98-128/Domain: EGF homology <EG1>
F;137-172/Domain: EGF homology <EG2>
F;197-456/Product: protein C heavy chain #status experimental <HCH>
F;197-210/Domain: activation peptide #status experimental <APT>
F;211-440/Domain: trypsin homology <TRY>
F;45,46,53,55,58,59,62,64,65,68,74/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;119-128,137-148,144-157,159-172,180-318,237-253,368-382,393-421/Disulfide bonds: #stat
F;136,289,350/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;252,298,397/Active site: His, Asp, Ser #status predicted
F;366/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             61.9%;  Score 122;  DB 1;  Length 456;
  Best Local Similarity   50.0%;  Pred. No. 1.2e-11;
  Matches   21;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
           |||||  || |:: | |   :|:|  |: ||:: :|||:||||
Db      40 ANSFLEELRPGNVERECSEEVCEFEEAREIFQNTEDTMAFWS 81


RESULT 5
EXRT
coagulation factor Xa (EC 3.4.21.6) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Jan-1995 #sequence_revision 07-Feb-1997 #text_change 08-Dec-2000
C;Accession: S49075; JC4670; PS0191; PS0190; I62745
R;Stanton, C.; Ross, P.; Hutson, S.; Wallin, R.
Thromb. Res. 80, 63-73, 1995
A;Title: Evidence for competition between vitamin K-dependent clotting factors for intra
A;Reference number: A58498; MUID:96093366; PMID:8578539
A;Accession: S49075
A;Molecule type: mRNA
A;Residues: 1-482 <STA1>
A;Cross-references: EMBL:X79807; NID:g506600; PIDN:CAA56202.1; PID:g506601
A;Note: submitted to the EMBL Data Library, June 1994
A;Note: neither the complete nucleic acid sequence nor the complete translation are s
R;Stanton, C.; Ross, R.P.; Hutson, S.; Wallin, R.
Gene 169, 269-273, 1996
A;Title: Processing and expression of rat and human clotting factor-X-encoding cDNAs.
A;Reference number: JC4670; MUID:96194815; PMID:8647460
A;Accession: JC4670
A;Molecule type: mRNA
A;Residues: 1-482 <STA2>
A;Cross-references: EMBL:X79807; NID:g506600; PIDN:CAA56202.1; PID:g506601
A;Experimental source: Cos-1 cell
R;Enjyoji, K.; Miyazaki, K.; Kato, H.
J. Biochem. 109, 890-898, 1991
A;Title: Characterization of rat factors X and Xa: demonstration of factor Xa in rat
A;Reference number: PS0190; MUID:92041742; PMID:1718949
A;Accession: PS0191
A;Molecule type: protein
A;Residues: 41-58,'X',60-65 <ENJ1>
A;Accession: PS0190
A;Molecule type: protein
A;Residues: 183-186,'X',188-207 <ENJ2>
R;Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
Eur. J. Haematol. 52, 162-168, 1994
A;Title: Analysis of the partial nucleotide sequences and deduced primary structures
A;Reference number: I46196; MUID:94222160; PMID:8168596
A;Accession: I62745
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 295-383,'G',385-455 <MUR>
A;Cross-references: GB:D21215; NID:g415309; PIDN:BAA04756.1; PID:g455396
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-179/Product: coagulation factor X light chain #status predicted <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-482/Product: coagulation factor X heavy chain #status predicted <HCH>
F;183-231/Domain: activation peptide #status predicted <APT>
F;232-482/Product: coagulation factor Xa heavy chain #status predicted <ACT>
F;232-460/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,79/Modified site: gamma-carboxyglutamic acid (Glu) #s
F;57-62,90-101,95-110,112-121,129-140,136-149,151-164,172-340,238-243,259-275,388-402
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;187/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;208/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;231-232/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #s
F;274,320,417/Active site: His, Asp, Ser #status predicted

  Query Match             58.4%;  Score 115;  DB 1;  Length 482;
  Best Local Similarity   43.2%;  Pred. No. 1.7e-10;
  Matches   19;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||||   :::|:| | |:  || |  |: :||| : |  ||:|:
Db      41 ANSFFEEIKKGNLERECVEEICSFEEAREVFEDNEKTTEFWNKY 84


RESULT 6
EXBO
coagulation factor Xa (EC 3.4.21.6) precursor - bovine
N;Alternate names: Stuart factor
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
```

C;Accession: A22867; A14997; A12030; A34412; S39414; A00925
R;Fung, M.R.; Campbell, R.M.; MacGillivray, T.A.
Nucleic Acids Res. 12, 4481-4492, 1984
A;Title: Blood coagulation factor X mRNA encodes a single polypeptide chain containing a
A;Reference number: A22867; MUID:84247315; PMID:6330671
A;Accession: A22867
A;Molecule type: mRNA
A;Residues: 1-487 <FUN>
A;Cross-references: GB:X00673; NID:g192; PIDN:CAA25286.1; PID:g193
R;Enfield, D.L.; Ericsson, L.H.; Fujikawa, K.; Walsh, K.A.; Neurath, H.; Titani, K.
Biochemistry 19, 659-667, 1980
A;Title: Amino acid sequence of the light chain of bovine factor X-1 (Stuart factor).
A;Reference number: A14997; MUID:80130563; PMID:6766735
A;Accession: A14997
A;Molecule type: protein
A;Residues: 41-102,'N',104-180 <ENF>
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood cc
A;Reference number: A20274; MUID:83308813; PMID:6688526
A;Contents: annotation; revision to residue 103
R;Titani, K.; Fujikawa, K.; Enfield, D.L.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.
Proc. Natl. Acad. Sci. U.S.A. 72, 3082-3086, 1975
A;Title: Bovine factor X-1 (Stuart factor): amino-acid sequence of heavy chain.
A;Reference number: A12030; MUID:76053069; PMID:1059093
A;Accession: A12030
A;Molecule type: protein
A;Residues: 183-292,294-295,'GDE',299-334,336-348,'AE',351-354,356-441,'GKFG',446-492 <T
A;Note: carbohydrate binding sites and disulfide bonds were determined
R;Persson, E.; Selander, M.; Linse, S.; Drakenberg, T.; Oehlin, A.K.; Stenflo, J.
J. Biol. Chem. 264, 16897-16904, 1989
A;Title: Calcium binding to the isolated beta-hydroxyaspartic acid-containing epidermal
A;Reference number: A34412; MUID:89380326; PMID:2789221
A;Accession: A34412
A;Molecule type: protein
A;Residues: 85-126 <PER>
A;Note: beta-hydroxyaspartic acid site
R;Inoue, K.; Morita, T.
Eur. J. Biochem. 218, 153-163, 1993
A;Title: Identification of O-linked oligosaccharide chains in the activation peptides of
A;Reference number: S39414; MUID:94062825; PMID:8243461
A;Accession: S39414
A;Molecule type: protein
A;Residues: 183-196;199-209;216-233 <INO>
A;Note: carbohydrate binding sites
R;Titani, K.; Hermodson, M.A.; Fujikawa, K.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.; D
Biochemistry 11, 4899-4903, 1972
A;Title: Bovine factor X-1a (activated Stuart factor). Evidence of homology with mammali
A;Reference number: A12453; MUID:73053314; PMID:4264286
A;Contents: annotation; active site
R;Fujikawa, K.; Titani, K.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 72, 3359-3363, 1975
A;Title: Activation of bovine factor X (Stuart factor): conversion of factor Xaalpha to
A;Reference number: A13504; MUID:76053121; PMID:1059122
A;Contents: annotation; activation
R;Sugo, T.; Bjork, I.; Holmgren, A.; Stenflo, J.
J. Biol. Chem. 259, 5705-5710, 1984
A;Title: Calcium-binding properties of bovine factor X lacking the gamma-carboxyglutamic
A;Reference number: A38024; MUID:84185716; PMID:6546930
A;Contents: annotation; calcium binding
R;Morita, T.; Jackson, C.M.
J. Biol. Chem. 261, 4008-4014, 1986
A;Reference number: A38025; MUID:86140210; PMID:3949800
A;Contents: annotation; sulfate binding
C;Comment: Factor Xa converts prothrombin to thrombin during blood clotting.
C;Comment: The two chains are formed from a single-chain precursor by the excision of tw
C;Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway), c
activation.
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with strc
C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamir
C;Genetics:
A;Gene: F10
A;Map position: 13q34

C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-180/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-492/Product: coagulation factor X heavy chain #status experimental <HCH>
F;183-233/Domain: activation peptide #status experimental <APT>
F;234-492/Product: coagulation factor Xa heavy chain #status experimental <AHC>
F;234-461/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu)
F;57-62,90-101,95-110,112-121,129-140,136-149,151-164,172-341/Disulfide bonds: #statu
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;200/Binding site: sulfate (Tyr) (covalent) (partial) #status experimental
F;208,485/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;233-234/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #s
F;240-245,260-276,389-403,414-442/Disulfide bonds: #status experimental
F;275,321,418/Active site: His, Asp, Ser #status predicted

Query Match 57.9%; Score 114; DB 1; Length 492;
Best Local Similarity 45.5%; Pred. No. 2.5e-10;
Matches 20; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

```
Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  ::||:| | |:   |    |: :||| : |  ||||:
Db      41 ANSFLEEVKQGNLERECLEEACSLEEAREVFEDAEQTDEFWSKY 84
```

RESULT 7
EXHU
coagulation factor Xa (EC 3.4.21.6) precursor [validated] - human
N;Alternate names: Stuart factor
C;Species: Homo sapiens (man)
C;Date: 15-Nov-1984 #sequence_revision 02-May-1994 #text_change 08-Dec-2000
C;Accession: A24478; JQ0917; A42485; A25853; A22208; A21284; A20362; S39415; I54051;
R;Leytus, S.P.; Foster, D.C.; Kurachi, K.; Davie, E.W.
Biochemistry 25, 5098-5102, 1986
A;Title: Gene for human Factor X: a blood coagulation factor whose gene organization
A;Reference number: A24478; MUID:87026600; PMID:3768336
A;Accession: A24478
A;Molecule type: DNA
A;Residues: 1-488 <LEY>
A;Cross-references: GB:L29433; GB:M14327; NID:g459809; PIDN:AAA52764.1; PID:g182831
R;Messier, T.L.; Pittman, D.D.; Long, G.L.; Kaufman, R.J.; Church, W.R.
Gene 99, 291-294, 1991
A;Title: Cloning and expression in COS-1 cells of a full-length cDNA encoding human c
A;Reference number: JQ0917; MUID:91216473; PMID:1902434
A;Accession: JQ0917
A;Molecule type: mRNA
A;Residues: 1-488 <MES>
A;Cross-references: GB:M57285; NID:g182389; PIDN:AAA52421.1; PID:g182390
R;Miao, C.H.; Leytus, S.P.; Chung, D.W.; Davie, E.W.
J. Biol. Chem. 267, 7395-7401, 1992
A;Title: Liver-specific expression of the gene coding for human factor X, a blood coa
A;Reference number: A42485; MUID:92218390; PMID:1313796
A;Accession: A42485
A;Molecule type: DNA
A;Residues: 1-15 <MIA>
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:93780, NCBIP:93787)
R;Kaul, R.K.; Hildebrand, B.; Roberts, S.; Jagadeeswaran, P.
Gene 41, 311-314, 1986
A;Title: Isolation and characterization of human blood-coagulation factor X cDNA.
A;Reference number: A25853; MUID:86221713; PMID:3011603
A;Accession: A25853
A;Molecule type: mRNA
A;Residues: 19-284,'E',289-488 <KAU>

```
A;Cross-references: GB:M22613; NID:g180335; PIDN:AAA51984.1; PID:g180336
R;Fung, M.R.; Hay, C.W.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 3591-3595, 1985
A;Title: Characterization of an almost full-length cDNA coding for human blood coagulati
A;Reference number: A22208; MUID:85216545; PMID:2582420
A;Accession: A22208
A;Molecule type: mRNA
A;Residues: 13-441,'S',443-488 <FUN>
A;Cross-references: GB:K03194; NID:g182840; PIDN:AAA52490.1; PID:g182841
R;Leytus, S.P.; Chung, D.W.; Kisiel, W.; Kurachi, K.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 3699-3702, 1984
A;Title: Characterization of a cDNA coding for human factor X.
A;Reference number: A21284; MUID:84222026; PMID:6587384
A;Accession: A21284
A;Molecule type: mRNA
A;Residues: 13-284,'E',289-488 <LE2>
A;Cross-references: GB:K01886
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.; Sasagawa, T.; Howald, W.N.; Kwa, E.Y.; Weins
Biochemistry 22, 2875-2884, 1983
A;Title: Complete amino acid sequence of the light chain of human blood coagulation fact
A;Reference number: A20362; MUID:83257207; PMID:6871167
A;Accession: A20362
A;Molecule type: protein
A;Residues: 41-179 <MCM>
R;Inoue, K.; Morita, T.
Eur. J. Biochem. 218, 153-163, 1993
A;Title: Identification of O-linked oligosaccharide chains in the activation peptides of
A;Reference number: S39414; MUID:94062825; PMID:8243461
A;Accession: S39415
A;Molecule type: protein
A;Residues: 183-234 <INO>
A;Note: glycosylation sites
A;Note: identification and characterization of beta-hydroxyaspartic acid
R;Jagadeeswaran, P.; Reddy, S.V.; Rao, K.J.; Hamsabhushanam, K.; Lyman, G.
Gene 84, 517-519, 1989
A;Title: Cloning and characterization of the 5' end (exon 1) of the gene encoding human
A;Reference number: I54051; MUID:90128299; PMID:2612918
A;Accession: I54051
A;Status: translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-23 <RES>
A;Cross-references: GB:M33297; NID:g183860; PIDN:AAA52636.1; PID:g553330
R;Padmanabhan, K.; Padmanabhan, K.P.; Tulinsky, A.; Park, C.H.; Bode, W.; Huber, R.; Bla
J. Mol. Biol. 232, 947-966, 1993
A;Title: Structure of human des(1-45) factor Xa at 2.2 angstroms resolution.
A;Reference number: A49458; MUID:93360277; PMID:8355279
A;Contents: annotation; X-ray crystallography, 2.2 angstroms
C;Comment: The two chains held together by one disulfide bond are formed from a single-c
C;Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway) or
C;Genetics:
A;Gene: GDB:F10
A;Cross-references: GDB:119890; OMIM:227600
A;Map position: 13q34-13q34
A;Introns: 24/1; 77/3; 86/1; 124/1; 150/3; 249/3; 289/1
A;Note: deficiency of this factor causes Stuart disease
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-179/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-164/Domain: EGF homology <EG2>
F;183-488/Product: coagulation factor X heavy chain #status experimental <HCH>
F;183-234/Domain: activation peptide #status experimental <APT>
F;235-488/Product: coagulation factor Xa heavy chain #status experimental <ACT>
F;235-462/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,79/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;57-62/Disulfide bonds: #status predicted
F;90-101,95-110,112-121,129-140,136-149,151-164,172-342,241-246,261-277,390-404,415-443/
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F;199,211/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;221,231/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;234-235/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #s
F;276,322,419/Active site: His, Asp, Ser #status experimental

  Query Match             55.8%;  Score 110;  DB 1;  Length 488;
  Best Local Similarity   43.2%;  Pred. No. 1.1e-09;
  Matches   19;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  :::| | | |:   | :  |: :||| | |  ||:|:
Db      41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKY 84


RESULT 8
I46932
coagulation factor VII - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 12-Feb-1999
C;Accession: I46932
R;Brothers, A.B.; Clarke, B.J.; Sheffield, W.P.; Blajchman, M.A.
Thromb. Res. 69, 231-238, 1993
A;Title: Complete nucleotide sequence of the cDNA encoding rabbit coagulation factor
A;Reference number: I46932; MUID:93190306; PMID:8383365
A;Accession: I46932
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-443 <BRO>
A;Cross-references: GB:S56300; NID:g266294; PID:g266295
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
F;24-83/Domain: Gla domain homology <GLA>
F;89-120/Domain: EGF homology <EG1>
F;130-166/Domain: EGF homology <EG2>
F;192-425/Domain: trypsin homology <TRY>

  Query Match             51.3%;  Score 101;  DB 2;  Length 443;
  Best Local Similarity   46.3%;  Pred. No. 2.8e-08;
  Matches   19;  Conservative    5;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
           |||||  || ||| | |   :| |  |: :|:  : |  ||
Db      40 ANSFLEELRPGSLERECKEELCSFEEAREVFQSTERTKQFW 80


RESULT 9
KFHU7
coagulation factor VIIa (EC 3.4.21.21) precursor [validated] - human
C;Species: Homo sapiens (man)
C;Date: 19-May-1989 #sequence_revision 19-May-1994 #text_change 08-Dec-2000
C;Accession: A28322; A23819; A31186; B31186; S63524
R;O'Hara, P.J.; Grant, F.J.; Haldeman, B.A.; Gray, C.L.; Insley, M.Y.; Hagen, F.S.; M
Proc. Natl. Acad. Sci. U.S.A. 84, 5158-5162, 1987
A;Title: Nucleotide sequence of the gene coding for human factor VII, a vitamin K-dep
A;Reference number: A28322; MUID:87260948; PMID:3037537
A;Accession: A28322
A;Molecule type: DNA
A;Residues: 1-466 <OHA>
A;Cross-references: GB:J02933; NID:g180333; PIDN:AAA51983.1; PID:g180334
R;Hagen, F.S.; Gray, C.L.; O'Hara, P.; Grant, F.J.; Saari, G.C.; Woodbury, R.G.; Hart
Proc. Natl. Acad. Sci. U.S.A. 83, 2412-2416, 1986
A;Title: Characterization of a cDNA coding for human factor VII.
A;Reference number: A23819; MUID:86205965; PMID:3486420
A;Accession: A23819
A;Molecule type: mRNA
A;Residues: 1-466 <HAG>
A;Cross-references: GB:M13232; NID:g182799; PIDN:AAA88040.1; PID:g182801
R;Thim, L.; Bjoern, S.; Christensen, M.; Nicolaisen, E.M.; Lund-Hansen, T.; Pedersen,
Biochemistry 27, 7785-7793, 1988
A;Title: Amino acid sequence and posttranslational modifications of human factor VII-
A;Reference number: A90539; MUID:89088153; PMID:3264725
A;Accession: A31186
A;Molecule type: protein
```

```
A;Residues: 61-212 <THI>
A;Accession: B31186
A;Molecule type: protein
A;Residues: 213-466 <TH2>
R;Bjoern, S.; Foster, D.C.; Thim, L.; Wiberg, F.C.; Christensen, M.; Komiyama, Y.; Peder
J. Biol. Chem. 266, 11051-11057, 1991
A;Title: Human plasma and recombinant factor VII. Characterization of O-glycosylations a
A;Reference number: A40529; MUID:91250411; PMID:1904059
A;Contents: annotation; carbohydrate binding sites
R;Persson, E.; Petersen, L.C.
Eur. J. Biochem. 234, 293-300, 1995
A;Title: Structurally and functionally distinct Ca(2+) binding sites in the gamma-carbox
A;Reference number: S63524; MUID:96096752; PMID:8529655
A;Accession: S63524
A;Molecule type: protein
A;Residues: 61-65;99-103;105-109;213-217;308-312 <PER>
C;Genetics:
A;Gene: GDB:F7
A;Cross-references: GDB:119897; OMIM:227500
A;Map position: 13q34-13q34
A;Introns: 22/1; 44/1; 97/3; 106/1; 144/1; 191/1; 227/3; 269/1
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor X in the prese
oagulation factor IX in the presence of calcium and tissue factor
A;Pathway: blood coagulation extrinsic pathway
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-60/Domain: propeptide #status predicted <PRO>
F;45-104/Domain: Gla domain homology <GLA>
F;61-212/Product: coagulation factor VIIa light chain #status experimental <MA1>
F;110-141/Domain: EGF homology <EG1>
F;151-187/Domain: EGF homology <EG2>
F;213-466/Product: coagulation factor VIIa heavy chain #status experimental <MA2>
F;213-447/Domain: trypsin homology <TRY>
F;66,67,74,76,79,80,85,86,89,95/Modified site: gamma-carboxyglutamic acid (Glu) #status
F;77-82,110-121,115-130,132-141,151-162,158-172,174-187,195-322,219-224,238-254,370-389,
F;112,120/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;123/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status absent
F;205,382/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;212-213/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;253,302,404/Active site: His, Asp, Ser #status predicted
F;350-351/Cleavage site: Arg-Gly (coagulation factor Xa) #status predicted

  Query Match             50.3%;  Score 99;  DB 1;  Length 466;
  Best Local Similarity   48.8%;  Pred. No. 6.1e-08;
  Matches   20;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
           ||:||  || ||| | |    | |  |: ||:| : |  ||
Db      61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 10
S10511
thrombin (EC 3.4.21.5) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 03-May-2002
C;Accession: S10511; A60576; B42696
R;Dihanich, M.; Monard, D.
Nucleic Acids Res. 18, 4251, 1990
A;Title: cDNA sequence of rat prothrombin.
A;Reference number: S10511; MUID:90332426; PMID:2377469
A;Accession: S10511
A;Molecule type: mRNA
A;Residues: 1-617 <DIH>
A;Cross-references: EMBL:X52835; NID:g56969; PIDN:CAA37017.1; PID:g56970
R;Henrikson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.
Endocrinology 126, 167-175, 1990
A;Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.
A;Reference number: A60576; MUID:90091942; PMID:2293980
A;Accession: A60576
A;Molecule type: protein
A;Residues: 44-58 <HEN>
A;Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A;Reference number: A42696; MUID:92212913; PMID:1557383
A;Accession: B42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 383-617,'E' <BAN>
A;Cross-references: GB:M81397
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; h
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-617/Product: prothrombin #status experimental <PMAT>
F;109-187/Domain: kringle homology <KR1>
F;215-292/Domain: kringle homology <KR2>
F;360-609/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;61-66,91-104,109-187,130-170,158-182,215-292,236-276,264-287,332-478,387-403,532-54
F;402,458,564/Active site: His, Asp, Ser #status predicted

  Query Match             43.9%;  Score 86.5;  DB 2;  Length 617;
  Best Local Similarity   42.2%;  Pred. No. 8.5e-06;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||| ||  ||:|:| | |:   | :  |    |   ||  ||:|:
Db      44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 11
A35827
thrombin (EC 3.4.21.5) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 03-May-2002
C;Accession: A35827; A42696; S12081
R;Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J
DNA Cell Biol. 9, 487-498, 1990
A;Title: Characterization of the cDNA coding for mouse prothrombin and localization o
A;Reference number: A35827; MUID:91025551; PMID:2222810
A;Accession: A35827
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-618 <DEG>
A;Cross-references: GB:X52308; NID:g53813; PIDN:CAA36548.1; PID:g53814
A;Experimental source: strain C57BL/6
A;Note: the data were obtained from females resulting from the cross of M. domesticus
R;Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A;Reference number: A42696; MUID:92212913; PMID:1557383
A;Accession: A42696
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 384-618,'E' <BAN>
A;Cross-references: GB:M81394
C;Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; h
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-43/Domain: propeptide #status predicted <PRO>
F;28-88/Domain: Gla domain homology <GLA>
F;44-618/Product: prothrombin B #status predicted <MAT>
F;109-187/Domain: kringle homology <KR1>
F;215-293/Domain: kringle homology <KR2>
F;361-610/Domain: trypsin homology <TRY>
F;50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F;61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-54
F;403,459,565/Active site: His, Asp, Ser #status predicted

  Query Match             43.9%;  Score 86.5;  DB 2;  Length 618;
```

```
  Best Local Similarity   42.2%;  Pred. No. 8.5e-06;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||| ||  ||:|:| | |:   | :  |    |   ||  ||:|:
Db      44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 12
EXCH
coagulation factor Xa (EC 3.4.21.6) precursor - chicken
N;Alternate names: virus-activating proteinase
C;Species: Gallus gallus (chicken)
C;Date: 12-Feb-1993 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C;Accession: S15838; S20380; S20381
R;Suzuki, H.; Harada, A.; Hayashi, Y.; Wada, K.; Asaka, J.; Gotoh, B.; Ogasawara, T.; Na
FEBS Lett. 283, 281-285, 1991
A;Title: Primary structure of the virus activating protease from chick embryo. Its ident
A;Reference number: S15838; MUID:91257322; PMID:2044767
A;Accession: S15838
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-475 <SUZ>
A;Cross-references: DDBJ:D00844; NID:g222869; PIDN:BAA00724.1; PID:g222870
R;Gotoh, B.; Yamauchi, F.; Ogasawara, T.; Nagai, Y.
FEBS Lett. 296, 274-278, 1992
A;Title: Isolation of factor Xa from chick embryo as the amniotic endoprotease responsib
A;Reference number: S20380; MUID:92164779; PMID:1537403
A;Accession: S20380
A;Molecule type: protein
A;Residues: 41-55 <GO2>
A;Accession: S20381
A;Molecule type: protein
A;Residues: 241-246,'X',248-251,'X',253-261 <GOT>
C;Function:
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr
A;Pathway: blood coagulation
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-40/Domain: propeptide #status predicted <PRO>
F;25-84/Domain: Gla domain homology <GLA>
F;41-185/Product: coagulation factor X light chain #status experimental <LCH>
F;90-121/Domain: EGF homology <EG1>
F;129-167/Domain: EGF homology <EG2>
F;186-475/Product: coagulation factor X heavy chain #status predicted <HCH>
F;186-240/Domain: activation peptide #status predicted <APT>
F;241-475/Product: coagulation factor Xa heavy chain #status experimental <AHC>
F;241-468/Domain: trypsin homology <TRY>
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #s
F;57-62,90-101,95-110,112-121,129-140,136-152,154-167,175-348,247-252,267-283,396-410,42
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;196,207,228,285/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;282,328,425/Active site: His, Asp, Ser #status predicted

  Query Match             43.7%;  Score 86;  DB 1;  Length 475;
  Best Local Similarity   40.5%;  Pred. No. 7.8e-06;
  Matches   17;  Conservative    7;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
           |||||  ::||:: | |    |    |:  ||| : |  ||:
Db      41 ANSFLEEMKQGNIERECNEERCSKEEAREAFEDNEKTEEFWN 82


RESULT 13
KFBO7
coagulation factor VIIa (EC 3.4.21.21) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-May-1990 #sequence_revision 23-Mar-1995 #text_change 16-Jul-1999
C;Accession: A31979; C20274
R;Takeya, H.; Kawabata, S.; Nakagawa, K.; Yamamichi, Y.; Miyata, T.; Iwanaga, S.
J. Biol. Chem. 263, 14868-14877, 1988
A;Title: Bovine factor VII. Its purification and complete amino acid sequence.
A;Reference number: A31979; MUID:89008362; PMID:3049594
A;Accession: A31979
A;Molecule type: protein
A;Residues: 1-407 <TAK>
R;McMullen, B.A.; Fujikawa, K.; Kisiel, W.
Biochem. Biophys. Res. Commun. 115, 8-14, 1983
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood
A;Reference number: A20274; MUID:83308813; PMID:6688526
A;Accession: C20274
A;Molecule type: protein
A;Residues: 58-62,'X',64-68 <MCM>
A;Note: the residue designated 'X' was determined to be hydroxyaspartic acid
R;Hase, S.; Kawabata, S.; Nishimura, H.; Takeya, H.; Sueyoshi, T.; Miyata, T.; Iwanag
J. Biochem. 104, 867-868, 1988
A;Title: A new trisaccharide sugar chain linked to a serine residue in bovine blood c
A;Reference number: A44556; MUID:89213999; PMID:3149637
A;Contents: annotation
A;Note: structure and location of covalently bound carbohydrate
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor X in the pr
gulation factor IX in the presence of calcium and tissue factor
A;Pathway: blood coagulation extrinsic pathway
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
F;1-152/Product: coagulation factor VIIa light chain #status experimental <MA1>
F;1-44/Domain: Gla domain homology (fragment) <GLA>
F;50-81/Domain: EGF homology <EG1>
F;91-127/Domain: EGF homology <EG2>
F;153-407/Product: coagulation factor VIIa heavy chain #status experimental <MA2>
F;153-387/Domain: trypsin homology <TRY>
F;6,7,14,16,19,20,25,26,29,34,35/Modified site: gamma-carboxyglutamic acid (Glu) #sta
F;17-22,50-61,55-70,72-81,91-102,98-112,114-127,135-262,159-164,178-194,310-329,340-3
F;52/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;63/Modified site: erythro-beta-hydroxyaspartic acid (Asp) (partial) #status experim
F;145,203/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;152-153/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;193,242,344/Active site: His, Asp, Ser #status predicted
F;290-291/Cleavage site: Arg-Gly (coagulation factor Xa) #status experimental

  Query Match             43.1%;  Score 85;  DB 1;  Length 407;
  Best Local Similarity   43.9%;  Pred. No. 9.7e-06;
  Matches   18;  Conservative    3;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
           || ||  |  ||| | |   :| | | || : : |  ||
Db       1 ANGFLEELLPGSLERECREELCSFEEAHEIFRNEERTRQFW 41


RESULT 14
S53434
plasma protein S precursor, vitamin K dependent - rhesus macaque (fragment)
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 16-Jul-1999
C;Accession: S53434
R;Greengard, J.S.; Fernandez, J.A.; Radtke, K.P.; Griffin, J.H.
Biochem. J. 305, 397-403, 1995
A;Title: Identification of candidate residues for interaction of protein S with C4b b
A;Reference number: S53433; MUID:95134217; PMID:7832752
A;Accession: S53434
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-642 <GRE>
A;Cross-references: EMBL:L31380
A;Experimental source: tissue type liver
A;Note: the source is designated as rhesus monkey
C;Genetics:
A;Gene: PROS
C;Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat
F;1-51/Domain: Gla domain homology (fragment) <GLA>
F;1-7/Domain: signal sequence (fragment) #status predicted <SIG>
F;8-642/Product: plasma protein S #status predicted <MAT>
F;87-120/Domain: EGF homology <EG1>
F;127-165/Domain: EGF homology <EG2>
```

F;171-207/Domain: EGF homology <EG3>
F;213-248/Domain: EGF homology <EG4>
F;281-633/Domain: sex hormone-binding globulin homology <SHB>
F;291-444/Domain: laminin G repeat homology <LGR>

Query Match 43.1%; Score 85; DB 2; Length 642;
Best Local Similarity 38.6%; Pred. No. 1.5e-05;
Matches 17; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||  |   :||:| | ||  :|:   |: :||:  :|  |: |:
Db      8 ANSMLEETKQGNLERECIEELCNKEEAREVFENDPETDYFYPKY 51
```

RESULT 15
KXHUS
plasma protein S precursor - human
N;Alternate names: vitamin K-dependent protein S
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1990 #sequence_revision 26-Jan-1996 #text_change 16-Jul-1999
C;Accession: A35610; A35611; A26157; A25891; A35612; A60903; S02424; S09519
R;Schmidel, D.K.; Tatro, A.V.; Phelps, L.G.; Tomczak, J.A.; Long, G.L.
Biochemistry 29, 7845-7852, 1990
A;Title: Organization of the human protein S genes.
A;Reference number: A35610; MUID:91084444; PMID:2148110
A;Accession: A35610
A;Molecule type: DNA
A;Residues: 1-676 <SCH>
A;Cross-references: GB:M57853; NID:g190547; PIDN:AAA60357.1; PID:g190549; GB:J02917
A;Note: the authors translated the codon TTT for residue 26 as Leu
R;Ploos van Amstel, H.K.; Reitsma, P.H.; van der Logt, C.P.E.; Bertina, R.M.
Biochemistry 29, 7853-7861, 1990
A;Title: Intron-exon organization of the active human protein S gene PSalpha and its pse
A;Reference number: A35611; MUID:91084445; PMID:2148111
A;Accession: A35611
A;Molecule type: DNA
A;Residues: 1-25 <PL3>
A;Cross-references: GB:J02918
R;Hoskins, J.; Norman, D.K.; Beckmann, R.J.; Long, G.L.
Proc. Natl. Acad. Sci. U.S.A. 84, 349-353, 1987
A;Title: Cloning and characterization of human liver cDNA encoding a protein S precursor
A;Reference number: A26157; MUID:87092407; PMID:3467362
A;Accession: A26157
A;Molecule type: mRNA
A;Residues: 1-10,'P',12-25,'L',27-676 <HOS>
A;Cross-references: GB:M15036; NID:g190288; PIDN:AAA36479.1; PID:g190289
R;Lundwall, A.; Dackowski, W.; Cohen, E.; Shaffer, M.; Mahr, A.; Dahlback, B.; Stenflo,
Proc. Natl. Acad. Sci. U.S.A. 83, 6716-6720, 1986
A;Title: Isolation and sequence of the cDNA for human protein S, a regulator of blood cc
A;Reference number: A25891; MUID:86313649; PMID:2944113
A;Accession: A25891
A;Molecule type: mRNA
A;Residues: 27-220,'L',222-262,'H',264-344,'Y',346-676 <LUN>
A;Cross-references: GB:M14338; NID:g190448; PIDN:AAA60181.1; PID:g190449
A;Note: part of this sequence, including the amino end of the mature protein, was determ
R;Edenbrandt, C.M.; Lundwall, A.; Wydro, R.; Stenflo, J.
Biochemistry 29, 7861-7868, 1990
A;Title: Molecular analysis of the gene for vitamin K dependent protein S and its pseudo
A;Reference number: A35612; MUID:91084446; PMID:2148112
A;Accession: A35612
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 284-676 <EDE>
A;Cross-references: GB:J02919
R;Ploos van Amstel, J.K.; van der Zanden, A.L.; Bakker, E.; Reitsma, P.H.; Bertina, R.M.
Thromb. Haemost. 58, 982-987, 1987
A;Title: Two genes homologous with human protein S cDNA are located on chromosome 3.
A;Reference number: A60903; MUID:88178564; PMID:2895503
A;Accession: A60903
A;Molecule type: mRNA
A;Residues: 351-676 <PLO>
R;Ploos van Amstel, H.K.; van der Zanden, A.L.; Reitsma, P.H.; Bertina, R.M.
FEBS Lett. 222, 186-190, 1987
A;Title: Human protein S cDNA encodes Phe-16 and Tyr 222 in consensus sequences for t
A;Reference number: S02424; MUID:88005138; PMID:2820795
A;Accession: S02424
A;Molecule type: mRNA
A;Residues: 1-676 <PL2>
A;Cross-references: EMBL:Y00692; NID:g36578; PIDN:CAA68687.1; PID:g36579
C;Genetics:
A;Gene: GDB:PROS1; PROS
A;Cross-references: GDB:120721; OMIM:176880
A;Map position: 3p11.1-3q11.2
A;Introns: 26/1; 78/3; 87/1; 116/1; 157/1; 201/1; 243/1; 283/3; 322/2; 385/3; 441/3;
C;Complex: in plasma forms a complex with C4b binding protein
C;Function:
A;Description: a cofactor for activated protein C (EC 3.4.21.69); thrombin cleavage d
C;Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat
C;Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coagulation; car
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-41/Domain: propeptide #status predicted <PRO>
F;26-85/Domain: Gla domain homology <GLA>
F;42-676/Product: plasma protein S #status predicted <MAT>
F;121-154/Domain: EGF homology <EG1>
F;161-199/Domain: EGF homology <EG2>
F;205-241/Domain: EGF homology <EG3>
F;247-282/Domain: EGF homology <EG4>
F;315-667/Domain: sex hormone-binding globulin homology <SHB>
F;325-478/Domain: laminin G repeat homology <LGR>
F;47,48,55,57,60,61,66,67,70,73,77/Modified site: gamma-carboxyglutamic acid (Glu) #s
F;58-63,88-113,121-134,126-143,145-154,161-175,171-184,186-199,205-217,212-226,228-24
F;111-112/Cleavage site: Arg-Ser (thrombin) #status predicted
F;136/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F;177,219,258/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;499,509,530/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.1%; Score 85; DB 1; Length 676;
Best Local Similarity 38.6%; Pred. No. 1.6e-05;
Matches 17; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||  |   :||:| | ||  :|:   |: :||:  :|  |: |:
Db     42 ANSLLEETKQGNLERECIEELCNKEEAREVFENDPETDYFYPKY 85
```

Search completed: May 15, 2003, 13:28:34
Job time : 19 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 15, 2003, 13:24:16 ; Search time 11 Seconds
(without alignments)
165.905 Million cell updates/sec

Title: SEQID1_MOD
Perfect score: 197
Sequence: 1 ANSFLXXLRQGSLXRXCIXX.....XXAKXIFEDVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	160	81.2	461	1	PRTC_HUMAN	P04070 homo sapien
2	140	71.1	461	1	PRTC_MOUSE	P33587 mus musculu
3	139	70.6	461	1	PRTC_RAT	P31394 rattus norv
4	138	70.1	458	1	PRTC_RABIT	Q28661 oryctolagus
5	123	62.4	459	1	PRTC_PIG	Q9glp2 sus scrofa
6	122	61.9	456	1	PRTC_BOVIN	P00745 bos taurus
7	114	57.9	492	1	FA10_BOVIN	P00743 bos taurus
8	110	55.8	488	1	FA10_HUMAN	P00742 homo sapien
9	107	54.3	231	1	TMG3_HUMAN	Q9bzd7 homo sapien
10	103	52.3	490	1	FA10_RABIT	O19045 oryctolagus
11	101	51.3	444	1	FA7_RABIT	P98139 oryctolagus
12	99	50.3	466	1	FA7_HUMAN	P08709 homo sapien
13	92	46.7	218	1	TMG1_HUMAN	O14668 homo sapien
14	86.5	43.9	617	1	THRB_RAT	P18292 rattus norv
15	86.5	43.9	618	1	THRB_MOUSE	P19221 mus musculu
16	86	43.7	475	1	FA10_CHICK	P25155 gallus gall
17	85	43.1	407	1	FA7_BOVIN	P22457 bos taurus
18	85	43.1	649	1	PRTS_MACMU	Q28520 macaca mula
19	85	43.1	676	1	PRTS_HUMAN	P07225 homo sapien
20	84.5	42.9	226	1	TMG4_HUMAN	Q9bzd6 homo sapien
21	84	42.6	622	1	THRB_HUMAN	P00734 homo sapien
22	82	41.6	376	1	FA10_TROCA	P81428 tropidechis
23	81	41.1	646	1	PRTS_RABIT	P98118 oryctolagus
24	80	40.6	446	1	FA7_MOUSE	P70375 mus musculu
25	80	40.6	452	1	FA9_CANFA	P19540 canis famil
26	80	40.6	459	1	FA9_MOUSE	P16294 mus musculu
27	80	40.6	461	1	FA9_HUMAN	P00740 homo sapien
28	80	40.6	675	1	PRTS_BOVIN	P07224 bos taurus
29	78	39.6	675	1	PRTS_RAT	P53813 rattus norv
30	73	37.1	416	1	FA9_BOVIN	P00741 bos taurus
31	72	36.5	625	1	THRB_BOVIN	P00735 bos taurus
32	71	36.0	675	1	PRTS_MOUSE	Q08761 mus musculu
33	69.5	35.3	396	1	PRTZ_BOVIN	P00744 bos taurus
34	65.5	33.2	400	1	PRTZ_HUMAN	P22891 homo sapien
35	65	33.0	202	1	TMG2_HUMAN	O14669 homo sapien
36	52	26.4	501	1	MKC1_CANAL	P43068 candida alb
37	50	25.4	363	1	ADK_TOXGO	Q9tvw2 toxoplasma
38	49	24.9	1363	1	VGR3_MOUSE	P35917 mus musculu
39	48	24.4	422	1	SPM1_SCHPO	Q92398 schizosacch
40	48	24.4	1235	1	CYA4_TRYBB	Q26721 trypanosoma
41	48	24.4	1298	1	VGR3_HUMAN	P35916 homo sapien
42	47	23.9	244	1	T2E5_ECOLI	P04390 escherichia
43	47	23.9	554	1	DHAB_SALTY	P37450 salmonella
44	47	23.9	1343	1	VGR2_RAT	O08775 rattus norv
45	47	23.9	1348	1	VGR2_COTJA	P52583 coturnix co

ALIGNMENTS

RESULT 1
PRTC_HUMAN
ID PRTC_HUMAN STANDARD; PRT; 461 AA.
AC P04070; Q16001; Q15190; Q15189;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE factor XIV).
GN PROC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270390; PubMed=2991887;
RA Foster D.C., Yoshitake S., Davie E.W.;
RT "The nucleotide sequence of the gene for human protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4673-4677(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85269639; PubMed=2991859;
RA Beckmann R.J., Schmidt R.J., Santerre R.F., Plutzky J., Crabtree G.R.,
RA Long G.L.;
RT "The structure and evolution of a 461 amino acid human protein C
RT precursor and its messenger RNA, based upon the DNA sequence of
RT cloned human liver cDNAs.";
RL Nucleic Acids Res. 13:5233-5247(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86120978; PubMed=3511471;
RA Plutzky J., Hoskins J.A., Long G.L., Crabtree G.R.;
RT "Evolution and organization of the human protein C gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:546-550(1986).
RN [4]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE OF 106-461 FROM N.A.
RX MEDLINE=84272714; PubMed=6589623;
RA Foster D.C., Davie E.W.;
RT "Characterization of a cDNA coding for human protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4766-4770(1984).
RN [6]
RP CARBOHYDRATE-LINKAGE SITE ASN-371.
RX MEDLINE=90293094; PubMed=1694179;
RA Miletich J.P., Broze G.J. Jr.;
RT "Beta protein C is not glycosylated at asparagine 329. The rate of
RT translation may influence the frequency of usage at asparagine-X-
RT cysteine sites.";
RL J. Biol. Chem. 265:11397-11404(1990).
RN [7]

RP HYDROXYLATION.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [8]
RP 3D-STRUCTURE MODELING OF 175-450.
RX MEDLINE=94272342; PubMed=8003977;
RA Fisher C.L., Greengard J.S., Griffin J.H.;
RT "Models of the serine protease domain of the human antithrombotic
RT plasma factor activated protein C and its zymogen.";
RL Protein Sci. 3:588-599(1994).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 84-461.
RX MEDLINE=97157472; PubMed=9003757;
RA Mather T., Oganessyan V., Hof P., Huber R., Foundling S., Esmon C.,
RA Bode W.;
RT "The 2.8 A crystal structure of Gla-domainless activated protein C.";
RL EMBO J. 15:6822-6831(1996).
RN [10]
RP REVIEW ON PROC VARIANTS.
RX MEDLINE=93190290; PubMed=8446940;
RA Reitsma P.H., Poort S.R., Bernardi F., Gandrille S., Long G.L.,
RA Sala N., Cooper D.N.;
RT "Protein C deficiency: a database of mutations. For the Protein C & S
RT Subcommittee of the Scientific and Standardization Committee of the
RT International Society on Thrombosis and Haemostasis.";
RL Thromb. Haemost. 69:77-84(1993).
RN [11]
RP VARIANT CYS-444.
RX MEDLINE=87204221; PubMed=2437584;
RA Romeo G., Hassan H.J., Staempfli S., Roncuzzi L., Cianetti L.,
RA Leonardi A., Vicente V., Mannucci P.M., Bertina R.M., Peschle C.,
RA Cortese R.;
RT "Hereditary thrombophilia: identification of nonsense and missense
RT mutations in the protein C gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2829-2832(1987).
RN [12]
RP VARIANT TRP-211 (LONDON-1).
RX MEDLINE=90098906; PubMed=2602169;
RA Grundy C.B., Chitolie A., Talbot S., Bevan D., Kakkar V.V.,
RA Cooper D.N.;
RT "Protein C London 1: recurrent mutation at Arg-169 (CGG-->TGG) in
RT the protein C gene causing thrombosis.";
RL Nucleic Acids Res. 17:10513-10513(1989).
RN [13]
RP VARIANT CYS-272.
RX MEDLINE=91329836; PubMed=1868249;
RA Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;
RT "The spectrum of genetic defects in a panel of 40 Dutch families with
RT symptomatic protein C deficiency type I: heterogeneity and founder
RT effects.";
RL Blood 78:890-894(1991).
RN [14]
RP VARIANTS ALA-62 (VERMONT-1) AND MET-76.
RX MEDLINE=92190481; PubMed=1347706;
RA Bovill E.G., Tomczak J.A., Grant B., Bhushan F., Pillemer E.,
RA Rainville I.R., Long G.L.;
RT "Protein CVermont: symptomatic type II protein C deficiency
RT associated with two GLA domain mutations.";
RL Blood 79:1456-1465(1992).
RN [15]
RP VARIANT ASP-418 (HONG KONG-2).
RX MEDLINE=92305321; PubMed=1611081;
RA Sugahara Y., Miura O., Yuen P., Aoki N.;
RT "Protein C deficiency Hong Kong 1 and 2: hereditary protein C
RT deficiency caused by two mutant alleles, a 5-nucleotide deletion and
RT a missense mutation.";
RL Blood 80:126-133(1992).
RN [16]
RP VARIANT LEU-289.
RX MEDLINE=92380660; PubMed=1511988;
RA Grundy C.B., Chisholm M., Kakkar V.V., Cooper D.N.;
RT "A novel homozygous missense mutation in the protein C (PROC) gene
RT causing recurrent venous thrombosis.";
RL Hum. Genet. 89:683-684(1992).
RN [17]
RP VARIANTS GLN-220 AND TRP-220.
RX MEDLINE=92380661; PubMed=1511989;
RA Grundy C.B., Schulman S., Tengborn L., Kakkar V.V., Cooper D.N.;
RT "Two different missense mutations at Arg 178 of the protein C (PROC)
RT gene causing recurrent venous thrombosis.";
RL Hum. Genet. 89:685-686(1992).
RN [18]
RP VARIANT GLN-220.
RX MEDLINE=93250852; PubMed=1301959;
RA Gandrille S., Vidaud M., Aiach M., Alhenc-Gelas M., Fischer A.M.,
RA Gouault-Heilman M., Toulon P., Fiessinger J.N., Goossens M.;
RT "Two novel mutations responsible for hereditary type I protein C
RT deficiency: characterization by denaturing gradient gel
RT electrophoresis.";
RL Hum. Mutat. 1:491-500(1992).
RN [19]
RP VARIANT SER-334.
RX MEDLINE=92276939; PubMed=1593215;
RA Yamamoto K., Matsushita T., Sugiura I., Takamatsu J., Iwasaki E.,
RA Wada H., Deguchi K., Shirakawa S., Saito H.;
RT "Homozygous protein C deficiency: identification of a novel missense
RT mutation that causes impaired secretion of the mutant protein C.";
RL J. Lab. Clin. Med. 119:682-689(1992).
RN [20]
RP VARIANTS TRP-38; CYS-42; HIS-42; GLN-271 AND ASN-294.
RX MEDLINE=93313192; PubMed=8324221;
RA Gandrille S., Alhenc-Gelas M., Gaussem P., Aillaud M.-F., Dupuy E.,
RA Juhan-Vague I., Aiach M.;
RT "Five novel mutations located in exons III and IX of the protein C
RT gene in patients presenting with defective protein C anticoagulant
RT activity.";
RL Blood 82:159-168(1993).
RN [21]
RP VARIANTS G-14; Q-211; Y-244; Q-253; L-321; C-328; I-385; T-388 AND
RP V-388.
RX MEDLINE=93271391; PubMed=8499565;
RA Poort S.R., Pabinger-Fasching I., Mannhalter C., Reitsma P.H.,
RA Bertina R.M.;
RT "Twelve novel and two recurrent mutations in 14 Austrian families
RT with hereditary protein C deficiency.";
RL Blood Coagul. Fibrinolysis 4:273-280(1993).
RN [22]
RP VARIANT TRP-57.
RX MEDLINE=93271396; PubMed=8499568;
RA Millar D.S., Grundy C.B., Bignell P., Moffat E.H., Martin R.,
RA Kakkar V.V., Cooper D.N.;
RT "A Gla domain mutation (Arg 15-->Trp) in the protein C (PROC) gene
RT causing type 2 protein C deficiency and recurrent venous
RT thrombosis.";
RL Blood Coagul. Fibrinolysis 4:345-347(1993).
RN [23]
RP VARIANTS R-145; L-210; W-211; T-243; L-321; M-340 AND Y-426.
RX MEDLINE=94122329; PubMed=8292730;
RA Tsay W., Greengard J.S., Montgomery R.R., McPherson R.A., Fucci J.C.,
RA Koerper M.A., Coughlin J., Griffin J.H.;
RT "Genetic mutations in ten unrelated American patients with
RT symptomatic type 1 protein C deficiency.";
RL Blood Coagul. Fibrinolysis 4:791-796(1993).
RN [24]
RP VARIANT SER-423.
RX MEDLINE=94001606; PubMed=8398832;
RA Marchetti G., Patracchini P., Gemmati D., Castaman G., Rodeghiero F.,
RA Wacey A., Cooper D.N., Tuddenham E.G., Bernardi F.;
RT "Symptomatic type II protein C deficiency caused by a missense
RT mutation (Gly 381-->Ser) in the substrate-binding pocket.";
RL Br. J. Haematol. 84:285-289(1993).
RN [25]
RP SEQUENCE OF 43-64 FROM N.A., AND VARIANT GLY-57 (YONAGO).

```
  Query Match             81.2%;  Score 160;  DB 1;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 9.1e-20;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  ||  || | ||  ||||  || ||::|||||||||||
Db      43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 2
PRTC_MOUSE
ID   PRTC_MOUSE     STANDARD;      PRT;   461 AA.
AC   P33587; O35498;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92316897; PubMed=1618739;
RA   Tada N., Sato M., Tsujimura A., Iwase R., Hashimoto-Gotoh T.;
RT   "Isolation and characterization of a mouse protein C cDNA.";
RL   J. Biochem. 111:491-495(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98152576; PubMed=9493582;
RA   Jalbert L.R., Rosen E.D., Lissens A., Carmeliet P., Collen D.,
RA   Castellino F.J.;
RT   "Nucleotide structure and characterization of the murine gene encoding
RT   anticoagulant protein C.";
RL   Thromb. Haemost. 79:310-316(1998).
RN   [3]
RP   SEQUENCE OF 274-434 FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=94318474; PubMed=8043441;
RA   Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT   "A comparative study of partial primary structures of the catalytic
RT   region of mammalian protein C.";
RL   Br. J. Haematol. 86:590-600(1994).
CC   -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC       REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC       IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC       INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC       BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC       TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
CC       REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC       STRONGLY PROMOTED BY THROMBOMODULIN.
CC   -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
CC   -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC       ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
CC       SITE IS NECESSARY FOR THE RECOGNITION OF THE
CC       THROMBIN-THROMBOMODULIN COMPLEX.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10445; BAA01235.1; -.
DR   EMBL; AF034569; AAC33795.1; -.
DR   EMBL; D43755; BAA07812.1; -.
DR   PIR; JX0210; JX0210.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   MGD; MGI:97771; Proc.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL       1     33       BY SIMILARITY.
FT   PROPEP      34     41       BY SIMILARITY.
FT   CHAIN       42    196       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN      199    461       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE    199    212       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE       212    213       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN      96    131       EGF-LIKE 1.
FT   DOMAIN     135    175       EGF-LIKE 2.
FT   DOMAIN     213    461       SERINE PROTEASE.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     48     48       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     57     57       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     67     67       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES     70     70       GAMMA-CARBOXYGLUTAMIC ACID
FT                               (BY SIMILARITY).
FT   MOD_RES    112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   253    253       CHARGE RELAY SYSTEM.
FT   ACT_SITE   299    299       CHARGE RELAY SYSTEM.
FT   ACT_SITE   402    402       CHARGE RELAY SYSTEM.
FT   DISULFID    58     63       BY SIMILARITY.
FT   DISULFID    91    110       BY SIMILARITY.
FT   DISULFID   100    105       BY SIMILARITY.
```

```
FT   DISULFID    104    119       BY SIMILARITY.
FT   DISULFID    121    130       BY SIMILARITY.
FT   DISULFID    139    150       BY SIMILARITY.
FT   DISULFID    146    159       BY SIMILARITY.
FT   DISULFID    161    174       BY SIMILARITY.
FT   DISULFID    182    319       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    238    254       BY SIMILARITY.
FT   DISULFID    373    387       BY SIMILARITY.
FT   DISULFID    398    426       BY SIMILARITY.
FT   CARBOHYD    214    214       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    290    290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    328    328       MISSING (IN REF. 2).
FT   CONFLICT    393    393       N -> D (IN REF. 2).
SQ   SEQUENCE   461 AA;  51945 MW;  53FAA0D85B194D6E CRC64;

  Query Match             71.1%;  Score 140;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 2.6e-16;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  :| ||| | |:  ||||  |: ||::|:|||||| |:
Db      42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
PRTC_RAT
ID   PRTC_RAT       STANDARD;      PRT;   461 AA.
AC   P31394;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=92329550; PubMed=1627650;
RA   Okafuji T., Maekawa K., Nawa K., Marumoto Y.;
RT   "The cDNA cloning and mRNA expression of rat protein C.";
RL   Biochim. Biophys. Acta 1131:329-332(1992).
CC   -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC       REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC       IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC       INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC       BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC       TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
CC       REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC       STRONGLY PROMOTED BY THROMBOMODULIN.
CC   -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
CC   -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC       ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
CC       SITE IS NECESSARY FOR THE RECOGNITION OF THE
CC       THROMBIN-THROMBOMODULIN COMPLEX.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X64336; CAA45617.1; -.
DR   PIR; S18994; S18994.
DR   PIR; S24312; S24312.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL        1     32       BY SIMILARITY.
FT   PROPEP       33     41       BY SIMILARITY.
FT   CHAIN        42    196       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN       199    461       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE     199    212       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE        212    213       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN       96    131       EGF-LIKE 1.
FT   DOMAIN      135    175       EGF-LIKE 2.
FT   DOMAIN      213    461       SERINE PROTEASE.
FT   MOD_RES      47     47       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      48     48       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      57     57       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      66     66       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      67     67       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES     112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE    254    254       CHARGE RELAY SYSTEM.
FT   ACT_SITE    300    300       CHARGE RELAY SYSTEM.
FT   ACT_SITE    402    402       CHARGE RELAY SYSTEM.
FT   DISULFID     58     63       BY SIMILARITY.
FT   DISULFID     91    110       BY SIMILARITY.
FT   DISULFID    100    105       BY SIMILARITY.
FT   DISULFID    104    119       BY SIMILARITY.
FT   DISULFID    121    130       BY SIMILARITY.
FT   DISULFID    139    150       BY SIMILARITY.
FT   DISULFID    146    159       BY SIMILARITY.
```

```
FT   DISULFID    161    174       BY SIMILARITY.
FT   DISULFID    182    320       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    239    255       BY SIMILARITY.
FT   DISULFID    373    387       BY SIMILARITY.
FT   DISULFID    398    426       BY SIMILARITY.
FT   CARBOHYD    215    215       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL)..
SQ   SEQUENCE   461 AA;  51912 MW;  8A4CF93664EDACD5 CRC64;

  Query Match             70.6%;  Score 139;  DB 1;  Length 461;
  Best Local Similarity   59.1%;  Pred. No. 3.9e-16;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  :| ||| | |:  ||||  |: ||::|:|||||| |:
Db      42 ANSFLEEVRAGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
PRTC_RABIT
ID   PRTC_RABIT     STANDARD;      PRT;   458 AA.
AC   Q28661;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV) (Fragment).
GN   PROC.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Shen L., He X., Dahlback B.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC       REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC       IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC       INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC       BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC       TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
CC       REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC       STRONGLY PROMOTED BY THROMBOMODULIN.
CC   -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
CC   -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC       ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
CC       SITE IS NECESSARY FOR THE RECOGNITION OF THE
CC       THROMBIN-THROMBOMODULIN COMPLEX.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U49933; AAA92956.1; -.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     27       BY SIMILARITY.
FT   PROPEP       28     36       BY SIMILARITY.
FT   CHAIN        37    458       VITAMIN K-DEPENDENT PROTEIN C.
FT   CHAIN        37    192       PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT   CHAIN       195    458       PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT   PEPTIDE     195    209       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE        209    210       CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT   DOMAIN       91    126       EGF-LIKE 1.
FT   DOMAIN      130    170       EGF-LIKE 2.
FT   DOMAIN      210    458       SERINE PROTEASE.
FT   MOD_RES      42     42       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      43     43       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      52     52       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      56     56       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      61     61       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      62     62       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID
FT                                (BY SIMILARITY).
FT   MOD_RES     107    107       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE    250    250       CHARGE RELAY SYSTEM.
FT   ACT_SITE    296    296       CHARGE RELAY SYSTEM.
FT   ACT_SITE    399    399       CHARGE RELAY SYSTEM.
FT   DISULFID     53     58       BY SIMILARITY.
FT   DISULFID     86    105       BY SIMILARITY.
FT   DISULFID     95    100       BY SIMILARITY.
FT   DISULFID     99    114       BY SIMILARITY.
FT   DISULFID    116    125       BY SIMILARITY.
FT   DISULFID    134    145       BY SIMILARITY.
FT   DISULFID    141    154       BY SIMILARITY.
FT   DISULFID    156    169       BY SIMILARITY.
FT   DISULFID    177    316       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    235    251       BY SIMILARITY.
FT   DISULFID    370    384       BY SIMILARITY.
FT   DISULFID    395    423       BY SIMILARITY.
FT   CARBOHYD    133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    287    287       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   458 AA;  51087 MW;  D75A5F990C8F29D7 CRC64;

  Query Match             70.1%;  Score 138;  DB 1;  Length 458;
  Best Local Similarity   59.1%;  Pred. No. 5.7e-16;
  Matches   26;  Conservative    4;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  ||  || | |:  :||   || ||: |||||||| |:
Db      37 ANSFLEELRPSSLERECVEEVCDLEEAKEIFQSVDDTLAFWYKY 80


RESULT 5
PRTC_PIG
ID   PRTC_PIG       STANDARD;      PRT;   459 AA.
AC   Q9GLP2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV).
GN   PROC.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=21121490; PubMed=11229814;
RA   Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,
RA   Kim H.K.W.;
RT   "Porcine factor V: cDNA cloning, gene mapping, three-dimensional
RT   protein modeling of membrane binding sites and comparative anatomy of
RT   domains.";
RL   Cell. Mol. Life Sci. 58:148-159(2001).
CC   -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC       regulates blood coagulation by inactivating factors Va and VIIIa
CC       in the presence of calcium ions and phospholipids.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC       into a light chain and a heavy chain held together by a disulfide
CC       bond. The enzyme is then activated by thrombin, which cleaves a
CC       tetradecapeptide from the amino end of the heavy chain; this
CC       reaction, which occurs at the surface of endothelial cells, is
CC       strongly promoted by thrombomodulin.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC       residues allows the modified protein to bind calcium.
CC   -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC       another site, beyond the GLA domain. This GLA-independent binding
CC       site is necessary for the recognition of the
CC       thrombin-thrombomodulin complex.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF191307; AAG28380.1; -.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00001; EGF_like; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   SIGNAL       1     18       BY SIMILARITY.
FT   PROPEP      19     41       BY SIMILARITY.
FT   CHAIN       42    459       VITAMIN K-DEPENDENT PROTEIN C.
FT   CHAIN       42    196       PROTEIN C LIGHT CHAIN (BY
FT                               SIMILARITY).
FT   CHAIN      199    459       PROTEIN C HEAVY CHAIN (BY
FT                               SIMILARITY).
FT   PEPTIDE    199    213       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   SITE       213    214       CLEAVAGE (BY THROMBIN) (BY
FT                               SIMILARITY).
FT   DOMAIN      96    131       EGF-LIKE 1.
FT   DOMAIN     135    175       EGF-LIKE 2.
FT   DOMAIN     214    459       SERINE PROTEASE.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     48     48       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     55     55       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     57     57       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     61     61       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     67     67       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES     70     70       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                               SIMILARITY).
FT   MOD_RES    112    112       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE   255    255       CHARGE RELAY SYSTEM.
FT   ACT_SITE   301    301       CHARGE RELAY SYSTEM.
FT   ACT_SITE   400    400       CHARGE RELAY SYSTEM.
FT   DISULFID    58     63       BY SIMILARITY.
FT   DISULFID    91    110       BY SIMILARITY.
FT   DISULFID   100    105       BY SIMILARITY.
FT   DISULFID   104    119       BY SIMILARITY.
FT   DISULFID   121    130       BY SIMILARITY.
FT   DISULFID   139    150       BY SIMILARITY.
FT   DISULFID   146    159       BY SIMILARITY.
FT   DISULFID   161    174       BY SIMILARITY.
FT   DISULFID   182    321       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   240    256       BY SIMILARITY.
FT   DISULFID   371    385       BY SIMILARITY.
FT   DISULFID   396    424       BY SIMILARITY.
FT   CARBOHYD   138    138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   459 AA;  51866 MW;  8541AAC14CC16D09 CRC64;

  Query Match             62.4%;  Score 123;  DB 1;  Length 459;
  Best Local Similarity   52.3%;  Pred. No. 2.2e-13;
  Matches   23;  Conservative    7;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  ||  || | |    |||  |: ||:: ::|:|||||:
Db     42 ANSFLEELRPSSLERECKEETCDFEEAREIFQNTENTMAFWSKY 85


RESULT 6
PRTC_BOVIN
ID   PRTC_BOVIN     STANDARD;      PRT;   456 AA.
AC   P00745;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE   (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
DE   factor XIV) (Fragment).
GN   PROC.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85014826; PubMed=6091100;
RA   Long G.L., Balagaje R.M., McGillivray R.T.A.;
RT   "Cloning and sequencing of liver cDNA coding for bovine protein C.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5653-5656(1984).
RN   [2]
RP   SEQUENCE OF 40-194.
RX   MEDLINE=83007325; PubMed=6896876;
RA   Fernlund P., Stenflo J.;
RT   "Amino acid sequence of the light chain of bovine protein C.";
RL   J. Biol. Chem. 257:12170-12179(1982).
RN   [3]
RP   REVISION TO 110.
RX   MEDLINE=83169769; PubMed=6572939;
RA   Drakenberg T., Fernlund P., Roepstorff P., Stenflo J.;
RT   "Beta-hydroxyaspartic acid in vitamin K-dependent protein C.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:1802-1806(1983).
RN   [4]
RP   SEQUENCE OF 197-456.
RX   MEDLINE=83007326; PubMed=6896877;
RA   Stenflo J., Fernlund P.;
RT   "Amino acid sequence of the heavy chain of bovine protein C.";
RL   J. Biol. Chem. 257:12180-12190(1982).
RN   [5]
RP   PROCESSING, AND CALCIUM-BINDING DATA.
RX   MEDLINE=83213513; PubMed=6304092;
RA   Esmon N.L., Debault L.E., Esmon C.T.;
RT   "Proteolytic formation and properties of gamma-carboxyglutamic acid-
RT   domainless protein C.";
RL   J. Biol. Chem. 258:5548-5553(1983).
RN   [6]
RP   PROCESSING, AND CALCIUM-BINDING DATA.
RX   MEDLINE=83213514; PubMed=6406503;
RA   Johnson A.E., Esmon N.L., Laue T.M., Esmon C.T.;
RT   "Structural changes required for activation of protein C are induced
RT   by Ca2+ binding to a high affinity site that does not contain gamma-
RT   carboxyglutamic acid.";
RL   J. Biol. Chem. 258:5554-5560(1983).
CC   -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC       REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC       IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC   -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC       and VIIIa.
CC   -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC       INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC       BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC       TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
CC       REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC       STRONGLY PROMOTED BY THROMBOMODULIN.
CC   -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
CC   -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC       ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
CC       SITE IS NECESSARY FOR THE RECOGNITION OF THE
CC       THROMBIN-THROMBOMODULIN COMPLEX.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02435; AAA30685.1; -.
DR   PIR; A00928; KXBO.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Glycoprotein; Serine protease;
KW   Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW   EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     29
FT   PROPEP       30     39
FT   CHAIN        40    194       PROTEIN C LIGHT CHAIN.
FT   CHAIN       197    456       PROTEIN C HEAVY CHAIN.
FT   PEPTIDE     197    210       ACTIVATION PEPTIDE.
FT   DOMAIN       94    129       EGF-LIKE 1.
FT   DOMAIN      133    173       EGF-LIKE 2.
FT   DOMAIN      211    456       SERINE PROTEASE.
FT   MOD_RES      45     45       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      53     53       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      62     62       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      68     68       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      74     74       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     110    110       HYDROXYLATION.
FT   ACT_SITE    252    252       CHARGE RELAY SYSTEM.
FT   ACT_SITE    298    298       CHARGE RELAY SYSTEM.
FT   ACT_SITE    397    397       CHARGE RELAY SYSTEM.
FT   DISULFID     56     61       BY SIMILARITY.
FT   DISULFID     89    108       BY SIMILARITY.
FT   DISULFID     98    103       BY SIMILARITY.
FT   DISULFID    102    117       BY SIMILARITY.
FT   DISULFID    119    128       BY SIMILARITY.
FT   DISULFID    137    148       BY SIMILARITY.
FT   DISULFID    144    157       BY SIMILARITY.
```

```
FT   DISULFID    159    172       BY SIMILARITY.
FT   DISULFID    180    318       INTERCHAIN.
FT   DISULFID    237    253
FT   DISULFID    368    382
FT   DISULFID    393    421
FT   CARBOHYD    136    136       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    289    289       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    366    366       N-LINKED (GLCNAC. . .).
FT   VARIANT      82     82       F -> K.
FT   CONFLICT    455    456       VP -> PV (IN REF. 4).
SQ   SEQUENCE   456 AA;  51407 MW;  CAAF6833F894C209 CRC64;

  Query Match             61.9%;  Score 122;  DB 1;  Length 456;
  Best Local Similarity   50.0%;  Pred. No. 3.3e-13;
  Matches   21;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
             |||||  || |:: | | :  :|:|  |: ||:: :||:|||| 
Db      40 ANSFLEELRPGNVERECSEEVCEFEEAREIFQNTEDTMAFWS 81


RESULT 7
FA10_BOVIN
ID   FA10_BOVIN     STANDARD;      PRT;   492 AA.
AC   P00743;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE OF 1-487 FROM N.A.
RX   MEDLINE=84247315; PubMed=6330671;
RA   Fung M.R., Campbell R.M., McGillivray R.T.A.;
RT   "Blood coagulation factor X mRNA encodes a single polypeptide chain
RT   containing a prepro leader sequence.";
RL   Nucleic Acids Res. 12:4481-4492(1984).
RN   [2]
RP   SEQUENCE OF 41-180.
RX   MEDLINE=80130563; PubMed=6766735;
RA   Enfield D.L., Ericsson L.H., Fujikawa K., Walsh K.A., Neurath H.,
RA   Titani K.;
RT   "Amino acid sequence of the light chain of bovine factor X1 (Stuart
RT   factor).";
RL   Biochemistry 19:659-667(1980).
RN   [3]
RP   REVISION TO 103.
RX   MEDLINE=83308813; PubMed=6688526;
RA   McMullen B.A., Fujikawa K., Kisiel W.;
RT   "The occurrence of beta-hydroxyaspartic acid in the vitamin
RT   K-dependent blood coagulation zymogens.";
RL   Biochem. Biophys. Res. Commun. 115:8-14(1983).
RN   [4]
RP   SEQUENCE OF 183-492, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RX   MEDLINE=76053069; PubMed=1059093;
RA   Titani K., Fujikawa K., Enfield D.L., Ericsson L.H., Walsh K.A.,
RA   Neurath H.;
RT   "Bovine factor X1 (Stuart factor): amino-acid sequence of heavey
RT   chain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 72:3082-3086(1975).
RN   [5]
RP   SEQUENCE OF 183-233, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=94062825; PubMed=8243461;
RA   Inoue K., Morita T.;
RT   "Identification of O-linked oligosaccharide chains in the activation
RT   peptides of blood coagulation factor X. The role of the carbohydrate
RT   moieties in the activation of factor X.";
RL   Eur. J. Biochem. 218:153-163(1993).
RN   [6]
RP   ACTIVE SITE.
RX   MEDLINE=73053314; PubMed=4264286;
RA   Titani K., Hermodson M.A., Fujikawa K., Ericsson L.H., Walsh K.A.,
RA   Neurath H., Davie E.W.;
RT   "Bovine factor X 1a (activated Stuart factor). Evidence of homology
RT   with mammalian serine proteases.";
RL   Biochemistry 11:4899-4903(1972).
RN   [7]
RP   PROCESSING.
RX   MEDLINE=76053121; PubMed=1059122;
RA   Fujikawa K., Titani K., Davie E.W.;
RT   "Activation of bovine factor X (Stuart factor): conversion of factor
RT   Xa-alpha to factor Xa-beta.";
RL   Proc. Natl. Acad. Sci. U.S.A. 72:3359-3363(1975).
RN   [8]
RP   CALCIUM-BINDING DATA.
RX   MEDLINE=84185716; PubMed=6546930;
RA   Sugo T., Bjoerk I., Holmgren A., Stenflo J.;
RT   "Calcium-binding properties of bovine factor X lacking the gamma-
RT   carboxyglutamic acid-containing region.";
RL   J. Biol. Chem. 259:5705-5710(1984).
RN   [9]
RP   SULFATION.
RX   MEDLINE=86140210; PubMed=3949800;
RA   Morita T., Jackson C.M.;
RT   "Localization of the structural difference between bovine blood
RT   coagulation factors X1 and X2 to tyrosine 18 in the activation
RT   peptide.";
RL   J. Biol. Chem. 261:4008-4014(1986).
RN   [10]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=91084483; PubMed=2261466;
RA   Selander M., Persson E., Stenflo J., Drakenberg T.;
RT   "1H NMR assignment and secondary structure of the Ca2(+)-free form of
RT   the amino-terminal epidermal growth factor like domain in coagulation
RT   factor X.";
RL   Biochemistry 29:8111-8118(1990).
RN   [11]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=92329412; PubMed=1627540;
RA   Ullner M., Selander M., Persson E., Stenflo J., Drakenberg T.,
RA   Teleman O.;
RT   "Three-dimensional structure of the apo form of the N-terminal
RT   EGF-like module of blood coagulation factor X as determined by NMR
RT   spectroscopy and simulated folding.";
RL   Biochemistry 31:5974-5983(1992).
RN   [12]
RP   STRUCTURE BY NMR OF 85-126.
RX   MEDLINE=92406922; PubMed=1527084;
RA   Selander-Sunnerhagen M., Ullner M., Persson E., Teleman O.,
RA   Stenflo J., Drakenberg T.;
RT   "How an epidermal growth factor (EGF)-like domain binds calcium. High
RT   resolution NMR structure of the calcium form of the NH2-terminal EGF-
RT   like domain in coagulation factor X.";
RL   J. Biol. Chem. 267:19642-19649(1992).
RN   [13]
RP   STRUCTURE BY NMR OF 41-126.
RX   MEDLINE=96387194; PubMed=8794734;
RA   Sunnerhagen M., Olah G.A., Stenflo J., Forsen S., Drakenberg T.,
RA   Trewhella J.;
RT   "The relative orientation of Gla and EGF domains in coagulation
RT   factor X is altered by Ca2+ binding to the first EGF domain. A
RT   combined NMR-small angle X-ray scattering study.";
RL   Biochemistry 35:11547-11559(1996).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR
CC       BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR
```

```
CC       MORE DISULFIDE BONDS.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC       CALCIUM.
CC   -!- PTM: N- AND O-GLYCOSYLATED.
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC   -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC       ANOTHER SITE, BEYOND THE GLA DOMAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X00673; CAA25286.1; -.
DR   PIR; A00925; EXBO.
DR   PDB; 1APO; 31-JAN-94.
DR   PDB; 1CCF; 31-MAY-94.
DR   PDB; 1WHE; 15-MAY-97.
DR   PDB; 1WHF; 15-MAY-97.
DR   MEROPS; S01.216; -.
DR   GlycoSuiteDB; P00743; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW   Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW   Signal; Zymogen; EGF-like domain; Repeat; Sulfation; 3D-structure.
FT   SIGNAL       1     23       POTENTIAL.
FT   PROPEP      24     40
FT   CHAIN       41    180       FACTOR X LIGHT CHAIN.
FT   CHAIN      183    492       FACTOR X HEAVY CHAIN.
FT   PROPEP     183    233       ACTIVATION PEPTIDE.
FT   CHAIN      234    492       ACTIVATED FACTOR XA, HEAVY CHAIN.
FT   PROPEP     476    492       MAY BE REMOVED BUT IS NOT NECESSARY FOR
FT                               ACTIVATION.
FT   DOMAIN      86    122       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     125    165       EGF-LIKE 2.
FT   DOMAIN     234    492       SERINE PROTEASE.
FT   ACT_SITE   275    275       CHARGE RELAY SYSTEM.
FT   ACT_SITE   321    321       CHARGE RELAY SYSTEM.
FT   ACT_SITE   418    418       CHARGE RELAY SYSTEM.
FT   MOD_RES     46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     47     47       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     54     54       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     56     56       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     65     65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     66     66       GAMMA-CARBOXYGLUTAMIC ACID.

  Query Match             57.9%;  Score 114;  DB 1;  Length 492;
  Best Local Similarity   45.5%;  Pred. No. 8.6e-12;
  Matches   20;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  ::||:| | |:   |    |: :||| : |  ||||:
Db      41 ANSFLEEVKQGNLERECLEEACSLEEAREVFEDAEQTDEFWSKY 84


RESULT 8
FA10_HUMAN
ID   FA10_HUMAN     STANDARD;      PRT:   488 AA.
AC   P00742; Q14340;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91216473; PubMed=1902434;
RA   Messier T.L., Pittman D.D., Long G.L., Kaufman R.J., Church W.R.;
RT   "Cloning and expression in COS-1 cells of a full-length cDNA encoding
RT   human coagulation factor X.";
RL   Gene 99:291-294(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87026600; PubMed=3768336;
RA   Leytus S.P., Foster D.C., Kurachi K., Davie E.W.;
RT   "Gene for human factor X: a blood coagulation factor whose gene
RT   organization is essentially identical with that of factor IX and
RT   protein C.";
RL   Biochemistry 25:5098-5102(1986).
RN   [3]
RP   SEQUENCE OF 13-488 FROM N.A.
RX   MEDLINE=85216545; PubMed=2582420;
RA   Fung M.R., Hay C.W., McGillivray R.T.A.;
RT   "Characterization of an almost full-length cDNA coding for human
RT   blood coagulation factor X.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3591-3595(1985).
RN   [4]
RP   SEQUENCE OF 19-488 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=86221713; PubMed=3011603;
RA   Kaul R.K., Hildebrand B., Roberts S., Jagadeeswaran P.;
RT   "Isolation and characterization of human blood-coagulation factor X
RT   cDNA.";
RL   Gene 41:311-314(1986).
RN   [5]
RP   SEQUENCE OF 41-179.
RX   MEDLINE=83257207; PubMed=6871167;
RA   McMullen B.A., Fujikawa K., Kisiel W., Sasagawa T., Howald W.N.,
RA   Kwa E.Y., Weinstein B.;
RT   "Complete amino acid sequence of the light chain of human blood
RT   coagulation factor X: evidence for identification of residue 63 as
RT   beta-hydroxyaspartic acid.";
RL   Biochemistry 22:2875-2884(1983).
RN   [6]
RP   SEQUENCE OF 115-488 FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Liver;
RX   MEDLINE=84222026; PubMed=6587384;
RA   Leytus S.P., Chung D.W., Kisiel W., Kurachi K., Davie E.W.;
RT   "Characterization of a cDNA coding for human factor X.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 81:3699-3702(1984).
RN   [7]
RP   SEQUENCE OF 183-234, AND CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=94062825; PubMed=8243461;
RA   Inoue K., Morita T.;
RT   "Identification of O-linked oligosaccharide chains in the activation
RT   peptides of blood coagulation factor X. The role of the carbohydrate
RT   moieties in the activation of factor X.";
RL   Eur. J. Biochem. 218:153-163(1993).
RN   [8]
RP   SEQUENCE OF 1-23 FROM N.A.
RX   MEDLINE=90128299; PubMed=2612918;
RA   Jagadeeswaran P., Reddy S.V., Rao K.J., Hamsabhushanam K., Lyman G.;
RT   "Cloning and characterization of the 5' end (exon 1) of the gene
RT   encoding human factor X.";
RL   Gene 84:517-519(1989).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 86-179 AND 235-278.
RX   MEDLINE=93360277; PubMed=8355279;
RA   Padmanabhan K., Padmanabhan K.P., Tulinsky A., Park C.H., Bode W.,
RA   Huber R., Blankenship D.T., Cardin A.D., Kisiel W.;
RT   "Structure of human des(1-45) factor Xa at 2.2-A resolution.";
RL   J. Mol. Biol. 232:947-966(1993).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 86-179 AND 235-278.
RX   MEDLINE=98283982; PubMed=9618463;
RA   Kamata K., Kawamoto H., Honma T., Iwama T., Kim S.H.;
RT   "Structural basis for chemical inhibition of human blood coagulation
RT   factor Xa.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:6630-6635(1998).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR
CC       BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR
CC       MORE DISULFIDE BONDS.
CC   -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC       CALCIUM.
CC   -!- PTM: N- AND O-GLYCOSYLATED.
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K03194; AAA52490.1; -.
DR   EMBL; M57285; AAA52421.1; -.
DR   EMBL; L29433; AAA52764.1; -.
DR   EMBL; L00390; AAA52764.1; JOINED.
DR   EMBL; L00391; AAA52764.1; JOINED.
DR   EMBL; L00392; AAA52764.1; JOINED.
DR   EMBL; L00393; AAA52764.1; JOINED.
DR   EMBL; L00394; AAA52764.1; JOINED.
DR   EMBL; L00395; AAA52764.1; JOINED.
DR   EMBL; L00396; AAA52764.1; JOINED.
DR   EMBL; M22613; AAA51984.1; -.
DR   EMBL; K01886; AAA52486.1; -.
DR   EMBL; M33297; AAA52636.1; -.
DR   PIR; A00924; EXHU.
DR   PIR; A25853; A25853.
DR   PIR; A24478; A24478.
DR   PDB; 1HCG; 08-MAY-95.
DR   PDB; 1FAX; 29-OCT-97.
DR   PDB; 1FXY; 17-JUN-98.
DR   PDB; 1XKA; 23-MAR-99.
DR   PDB; 1XKB; 23-MAR-99.
DR   MEROPS; S01.216; -.
DR   GlycoSuiteDB; P00742; -.
DR   Genew; HGNC:3528; F10.
DR   MIM; 134530; -.
DR   MIM; 227600; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW   Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW   Signal; Zymogen; EGF-like domain; Repeat; 3D-structure.
FT   SIGNAL       1    31       POTENTIAL.
FT   PROPEP      32    40
FT   CHAIN       41   179       FACTOR X LIGHT CHAIN.
FT   CHAIN      183   488       FACTOR X HEAVY CHAIN.
FT   PROPEP     183   234       ACTIVATION PEPTIDE.
FT   CHAIN      235   488       ACTIVATED FACTOR XA, HEAVY CHAIN.
FT   DOMAIN      86   122       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     125   165       EGF-LIKE 2.
FT   DOMAIN     235   488       SERINE PROTEASE.
FT   MOD_RES     46    46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     47    47       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     54    54       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     56    56       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     59    59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     60    60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     65    65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     66    66       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     69    69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     72    72       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     79    79       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES    103   103       HYDROXYLATION.
FT   CARBOHYD   199   199       O-LINKED (GALNAC. . .).
FT   CARBOHYD   211   211       O-LINKED (GALNAC. . .).
FT   CARBOHYD   221   221       N-LINKED (GLCNAC. . .).
FT                              /FTId=CAR_000012.
FT   CARBOHYD   231   231       N-LINKED (GLCNAC. . .).
FT                              /FTId=CAR_000013.
FT   ACT_SITE   276   276       CHARGE RELAY SYSTEM.
FT   ACT_SITE   322   322       CHARGE RELAY SYSTEM.
FT   ACT_SITE   419   419       CHARGE RELAY SYSTEM.
FT   DISULFID    90   101
FT   DISULFID    95   110
FT   DISULFID   112   121
FT   DISULFID   129   140

  Query Match             55.8%;  Score 110;  DB 1;  Length 488;
```

```
  Best Local Similarity   43.2%;  Pred. No. 4.2e-11;
  Matches   19;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  :::| | | |:   | :  |: :||| | |  ||:|:
Db      41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKY 84


RESULT 9
TMG3_HUMAN
ID   TMG3_HUMAN     STANDARD;      PRT;   231 AA.
AC   Q9BZD7;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Transmembrane gamma-carboxyglutamic acid protein 3 precursor.
GN   TMG3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spinal cord;
RX   MEDLINE=21117044; PubMed=11171957;
RA   Kulman J.D., Harris J.E., Xie L., Davie E.W.;
RT   "Identification of two novel transmembrane gamma-carboxyglutamic acid
RT   proteins expressed broadly in fetal and adult tissues.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375(2001).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in brain, lung, kidney and heart.
CC   -!- PTM: Gla residues are produced after subsequent posttranslational
CC       modifications of glutamic acid by a vitamin K-dependent gamma-
CC       carboxylase.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF326350; AAK00955.1; -.
DR   HSSP; P00740; 1CFH.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Gamma-carboxyglutamic acid; Vitamin K; Transmembrane.
FT   PROPEP          1     19       POTENTIAL.
FT   CHAIN          20    231       TRANSMEMBRANE GAMMA-CARBOXYGLUTAMIC ACID
FT                                  PROTEIN 3.
FT   DOMAIN         20     78       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       79    101       POTENTIAL.
FT   DOMAIN        102    231       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         23     60       GLA-RICH.
SQ   SEQUENCE   231 AA;  25848 MW;  8A373E4848490D81 CRC64;

  Query Match             54.3%;  Score 107;  DB 1;  Length 231;
  Best Local Similarity   43.9%;  Pred. No. 5.9e-11;
  Matches   18;  Conservative    8;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
           || ||  ||||:: | |:  || :   | :||: : |: ||
Db      20 ANEFLEELRQGTIERECMEEICSYEEVKEVFENKEKTMEFW 60


RESULT 10
FA10_RABIT
ID   FA10_RABIT     STANDARD;      PRT;   490 AA.
AC   O19045;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN   F10.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97256311; PubMed=9101642;
RA   Pendurthi U.R., Anderson K.D., James H.L.;
RT   "Characterization of a full-length cDNA for rabbit factor X.";
RL   Thromb. Res. 85:503-514(1997).
CC   -!- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC       converts prothrombin to thrombin in the presence of factor Va,
CC       calcium and phospholipid during blood clotting.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC       Arg-|-Ile bonds in prothrombin to form thrombin.
CC   -!- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR
CC       BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR
CC       MORE DISULFIDE BONDS.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC       CALCIUM (BY SIMILARITY).
CC   -!- PTM: N- AND O-GLYCOSYLATED (BY SIMILARITY).
CC   -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC       INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)
CC       (BY SIMILARITY).
CC   -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC       ANOTHER SITE, BEYOND THE GLA DOMAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003200; AAB62542.1; -.
DR   HSSP; P00742; 1HCG.
DR   MEROPS; S01.216; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
```

```
KW   Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW   Signal; Zymogen; EGF-like domain; Repeat.
FT   SIGNAL        1     20       POTENTIAL.
FT   PROPEP       21     40       BY SIMILARITY.
FT   CHAIN        41    180       FACTOR X LIGHT CHAIN.
FT   CHAIN       184    490       FACTOR X HEAVY CHAIN.
FT   PROPEP      184    232       ACTIVATION PEPTIDE.
FT   CHAIN       233    490       ACTIVATED FACTOR XA, HEAVY CHAIN.
FT   DOMAIN       86    122       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      125    165       EGF-LIKE 2.
FT   DOMAIN      233    490       SERINE PROTEASE.
FT   MOD_RES      46     46       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      47     47       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      54     54       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      56     56       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      59     59       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      66     66       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      72     72       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      75     75       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      79     79       GAMMA-CARBOXYGLUTAMIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES     103    103       HYDROXYLATION (BY SIMILARITY).
FT   ACT_SITE    274    274       CHARGE RELAY SYSTEM.
FT   ACT_SITE    320    320       CHARGE RELAY SYSTEM.
FT   ACT_SITE    417    417       CHARGE RELAY SYSTEM.
FT   DISULFID     90    101       BY SIMILARITY.
FT   DISULFID     95    110       BY SIMILARITY.
FT   DISULFID    112    121       BY SIMILARITY.
FT   DISULFID    129    140       BY SIMILARITY.
FT   DISULFID    136    149       BY SIMILARITY.
FT   DISULFID    151    164       BY SIMILARITY.
FT   DISULFID    172    340       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    239    244       BY SIMILARITY.
FT   DISULFID    259    275       BY SIMILARITY.
FT   DISULFID    388    402       BY SIMILARITY.
FT   DISULFID    413    441       BY SIMILARITY.
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    205    205       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   490 AA;  53965 MW;  3A39FA85AF2A6D11 CRC64;

  Query Match             52.3%;  Score 103;  DB 1;  Length 490;
  Best Local Similarity   43.2%;  Pred. No. 6.8e-10;
  Matches   19;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  |::|:| | |:   | :  |  :||| : |  ||:|:
Db      41 ANSFLEELKKGNLERECMEENCSYEEALEVFEDREKTNEFWNKY 84


RESULT 11
FA7_RABIT
ID   FA7_RABIT      STANDARD;      PRT;   444 AA.
AC   P98139; P79224;
DT   01-FEB-1996 (Rel. 33, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin
DE   conversion accelerator).
GN   F7.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93190306; PubMed=8383365;
RA   Brothers A.B., Clarke B.J., Sheffield W.P., Blajchman M.A.;
RT   "Complete nucleotide sequence of the cDNA encoding rabbit coagulation
RT   factor VII.";
RL   Thromb. Res. Suppl. 69:231-238(1993).
RN   [2]
RP   REVISION TO 395.
RC   TISSUE=Liver;
RA   Ruiz S.R., Blajchman M.A., Clarke B.J.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CIRCULATES IN THE BLOOD IN A ZYMOGEN FORM. FACTOR VII IS
CC       CONVERTED TO FACTOR VIIA BY FACTOR XA, FACTOR XIIA, FACTOR IXA, OR
CC       THROMBIN BY MINOR PROTEOLYSIS. IN THE PRESENCE OF TISSUE FACTOR
CC       AND CALCIUM IONS, FACTOR VIIA THEN CONVERTS FACTOR X TO FACTOR XA
CC       BY LIMITED PROTEOLYSIS. FACTOR VIIA WILL ALSO CONVERT FACTOR IX TO
CC       FACTOR IXA IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
CC       form factor Xa.
CC   -!- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED
CC       BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC       CALCIUM (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U77477; AAB37326.1; -.
DR   HSSP; P08709; 1FAK.
DR   MEROPS; S01.215; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Blood coagulation; Zymogen; Glycoprotein;
```

```
KW   Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;
KW   EGF-like domain; Repeat; Signal; Hydroxylation.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     39       POTENTIAL.
FT   CHAIN        40    191       FACTOR VII LIGHT CHAIN.
FT   CHAIN       192    444       FACTOR VII HEAVY CHAIN.
FT   DOMAIN       45     74       GLA-RICH.
FT   DOMAIN       85    121       EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      126    167       EGF-LIKE 2.
FT   DOMAIN      192    444       SERINE PROTEASE.
FT   SITE        191    192       CLEAVAGE (BY FACTOR XA, FACTOR XIIA,
FT                                FACTOR IXA, OR THROMBIN) (BY SIMILARITY).
FT   ACT_SITE    232    232       BY SIMILARITY.
FT   ACT_SITE    281    281       BY SIMILARITY.
FT   ACT_SITE    383    383       BY SIMILARITY.
FT   BINDING     377    377       SUBSTRATE (BY SIMILARITY).
FT   DISULFID     56     61       BY SIMILARITY.
FT   DISULFID     89    100       BY SIMILARITY.
FT   DISULFID     94    109       BY SIMILARITY.
FT   DISULFID    111    120       BY SIMILARITY.
FT   DISULFID    130    141       BY SIMILARITY.
FT   DISULFID    137    151       BY SIMILARITY.
FT   DISULFID    153    166       BY SIMILARITY.
FT   DISULFID    174    301       BY SIMILARITY.
FT   DISULFID    198    203       BY SIMILARITY.
FT   DISULFID    217    233       BY SIMILARITY.
FT   DISULFID    349    368       BY SIMILARITY.
FT   DISULFID    379    407       BY SIMILARITY.
FT   MOD_RES      45     45       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      46     46       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      53     53       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      55     55       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      59     59       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      65     65       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      68     68       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      74     74       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES     102    102       HYDROXYLATION (BY SIMILARITY).
FT   CARBOHYD    211    211       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    306    306       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   444 AA;  49011 MW;  0481ABC4FE5427F8 CRC64;

  Query Match           51.3%;  Score 101;  DB 1;  Length 444;
  Best Local Similarity   46.3%;  Pred. No. 1.3e-09;
  Matches   19;  Conservative    5;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
             |||||  || ||| | |   :| |  |: :|:  : |  ||
Db      40 ANSFLEELRPGSLERECKEELCSFEEAREVFQSTERTKQFW 80


RESULT 12
FA7_HUMAN
ID   FA7_HUMAN      STANDARD;      PRT;   466 AA.
AC   P08709; Q14339;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin
DE   conversion accelerator) (Eptacog alfa).
GN   F7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=86205965; PubMed=3486420;
RA   Hagen F.S., Gray C.L., O'Hara P.J., Grant F.J., Saari G.C.,
RA   Woodbury R.G., Hart C.E., Insley M.Y., Kisiel W., Kurachi K.,
RA   Davie E.W.;
RT   "Characterization of a cDNA coding for human factor VII.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:2412-2416(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87260948; PubMed=3037537;
RA   O'Hara P.J., Grant F.J., Haldeman B.A., Gray C.L., Insley M.Y.,
RA   Hagen F.S., Murray M.J.;
RT   "Nucleotide sequence of the gene coding for human factor VII, a
RT   vitamin K-dependent protein participating in blood coagulation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:5158-5162(1987).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS THR-352; GLN-413 AND LYS-445.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 61-466, AND POST-TRANSLATIONAL MODIFICATIONS.
RX   MEDLINE=89088153; PubMed=3264725;
RA   Thim L., Bjoern S., Christensen M., Nicolaisen E.M., Lund-Hansen T.,
RA   Pedersen A.H., Hedner U.;
RT   "Amino acid sequence and posttranslational modifications of human
RT   factor VIIa from plasma and transfected baby hamster kidney cells.";
RL   Biochemistry 27:7785-7793(1988).
RN   [5]
RP   CARBOHYDRATE-LINKAGE SITES SER-112 AND SER-120.
RX   MEDLINE=91250411; PubMed=1904059;
RA   Bjoern S., Foster D.C., Thim L., Wiberg F.C., Christensen M.,
RA   Komiyama Y., Pedersen A.H., Kisiel W.;
RT   "Human plasma and recombinant factor VII. Characterization of O-
RT   glycosylations at serine residues 52 and 60 and effects of site-
RT   directed mutagenesis of serine 52 to alanine.";
RL   J. Biol. Chem. 266:11051-11057(1991).
RN   [6]
RP   STRUCTURE OF CARBOHYDRATE ON SER-112.
RX   MEDLINE=90062160; PubMed=2511201;
RA   Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T., Takao T.,
RA   Shimonishi Y., Iwanaga S.;
RT   "Identification of a disaccharide (Xyl-Glc) and a trisaccharide
RT   (Xyl2-Glc) O-glycosidically linked to a serine residue in the first
RT   epidermal growth factor-like domain of human factors VII and IX and
RT   protein Z and bovine protein Z.";
RL   J. Biol. Chem. 264:20320-20325(1989).
RN   [7]
RP   STRUCTURE OF CARBOHYDRATE ON SER-112.
RX   MEDLINE=91344709; PubMed=2129367;
RA   Iwanaga S., Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.;
RT   "A new trisaccharide sugar chain linked to a serine residue in the
RT   first EGF-like domain of clotting factors VII and IX and protein Z.";
RL   Adv. Exp. Med. Biol. 281:121-131(1990).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
RX   MEDLINE=96175641; PubMed=8598903;
RA   Banner D.W., D'Arcy A., Chene C., Winkler F.K., Guha A.,
RA   Konigsberg W.H., Nemreson Y., Kirchhofer D.;
RT   "The crystal structure of the complex of blood coagulation factor
RT   VIIa with soluble tissue factor.";
RL   Nature 380:41-46(1996).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
RX   MEDLINE=99126538; PubMed=9925787;
RA   Zhang E., St Charles R., Tulinsky A.;
RT   "Structure of extracellular tissue factor complexed with factor VIIa
RT   inhibited with a BPTI mutant.";
RL   J. Mol. Biol. 285:2089-2104(1999).
RN   [10]
RP   STRUCTURE BY NMR OF 105-145.
RX   MEDLINE=98367502; PubMed=9692950;
RA   Muranyi A., Finn B.E., Gippert G.P., Forsen S., Stenflo J.,
RA   Drakenberg T.;
RT   "Solution structure of the N-terminal EGF-like domain from human
RT   factor VII.";
RL   Biochemistry 37:10605-10615(1998).
```

```
RN   [11]
RP   VARIANT GLN-364.
RX   MEDLINE=91300046; PubMed=2070047;
RA   O'Brien D.P., Gale K.M., Anderson J.S., McVey J.H., Miller G.J.,
RA   Meade T.W., Tuddenham E.G.D.;
RT   "Purification and characterization of factor VII 304-Gln: a variant
RT   molecule with reduced activity isolated from a clinically unaffected
RT   male.";
RL   Blood 78:132-140(1991).
RN   [12]
RP   VARIANTS GLN-364 AND PHE-370.
RX   MEDLINE=92340074; PubMed=1634227;
RA   Marchetti G., Patracchini P., Gemmati D., Derosa V., Pinotti M.,
RA   Rodorigo G., Casonato A., Girolami A., Bernardi F.;
RT   "Detection of two missense mutations and characterization of a repeat
RT   polymorphism in the factor VII gene (F7).";
RL   Hum. Genet. 89:497-502(1992).
RN   [13]
RP   VARIANT TYR-238.
RX   MEDLINE=93372811; PubMed=8364544;
RA   Marchetti G., Ferrati M., Patracchini P., Redaelli R., Bernardi F.;
RT   "A missense mutation (178Cys-->Tyr) and two neutral dimorphisms
RT   (115His and 333Ser) in the human coagulation factor VII gene.";
RL   Hum. Mol. Genet. 2:1055-1056(1993).
RN   [14]
RP   VARIANTS.
RX   MEDLINE=94061028; PubMed=8242057;
RA   Takamiya O., Kemball-Cook G., Marin D.M.A., Cooper D.N.,
RA   von Felten A., Meili E., Hahn I., Prangnell D.R., Lumley H.,
RA   Tuddenham E.G.D., McVey J.H.;
RT   "Detection of missense mutations by single-strand conformational
RT   polymorphism (SSCP) analysis in five dysfunctional variants of
RT   coagulation factor VII.";
RL   Hum. Mol. Genet. 2:1355-1359(1993).
RN   [15]
RP   VARIANTS CHARLOTTE GLN-139 AND GLN-212.
RX   MEDLINE=94264305; PubMed=8204879;
RA   Chaing S., Clarke B., Sridhara S., Chu K., Friedman P., Vandusen W.,
RA   Roberts H.R., Blajchman M., Monroe D.M., High K.A.;
RT   "Severe factor VII deficiency caused by mutations abolishing the
RT   cleavage site for activation and altering binding to tissue factor.";
RL   Blood 83:3524-3535(1994).
RN   [16]
RP   VARIANT VAL-354.
RX   MEDLINE=95072589; PubMed=7981691;
RA   Bernardi F., Castaman G., Redaelli R., Pinotti M., Lunghi B.,
RA   Rodeghiero F., Marchetti G.;
RT   "Topologically equivalent mutations causing dysfunctional coagulation
RT   factors VII (294Ala-->Val) and X (334Ser-->Pro).";
RL   Hum. Mol. Genet. 3:1175-1177(1994).
RN   [17]
RP   VARIANT MIE HIS-307.
RX   MEDLINE=95064662; PubMed=7974346;
RA   Ohiwa M., Hayashi T., Wada H., Minamikawa K., Shirakawa S.,
RA   Suzuki K.;
RT   "Factor VII Mie: homozygous asymptomatic type I deficiency caused by
RT   an amino acid substitution of His (CAC) for Arg(247) (CGC) in the
RT   catalytic domain.";
RL   Thromb. Haemost. 71:773-777(1994).
RN   [18]
RP   VARIANT MET-419.
RX   MEDLINE=96247510; PubMed=8652821;
RA   Arbini A.A., Mannucci P.M., Bauer K.A.;
RT   "A Thr359Met mutation in factor VII of a patient with a hereditary
RT   deficiency causes defective secretion of the molecule.";
RL   Blood 87:5085-5094(1996).
RN   [19]
RP   VARIANTS W-283; K-325; V-358; Q-364; E-402 AND Q-413.
RX   MEDLINE=97001216; PubMed=8844208;
RA   Bernardi F., Castaman G., Pinotti M., Ferraresi P., di Iasio M.G.,
RA   Lunghi B., Rodeghiero F., Marchetti G.;
RT   "Mutation pattern in clinically asymptomatic coagulation factor VII
RT   deficiency.";
RL   Hum. Mutat. 8:108-115(1996).
RN   [20]
RP   VARIANT VAL-304.
RX   MEDLINE=97037613; PubMed=8883260;
RA   Tamary H., Fromovich Y., Shalmon L., Reich Z., Dym O., Lanir N.,
RA   Brenner B., Paz M., Luder A.S., Blau O., Korostishevsky M.,
RA   Zaizov R., Seligsohn U.;
RT   "Ala244Val is a common, probably ancient mutation causing factor VII
RT   deficiency in Moroccan and Iranian Jews.";
RL   Thromb. Haemost. 76:283-291(1996).
RN   [21]
RP   VARIANTS MALTA THR-194 AND VAL-304.
RX   MEDLINE=98112461; PubMed=9452082;
RA   Alshinawi C., Scerri C., Galdies R., Aquilina A., Felice A.E.;
RT   "Two new missense mutations (P134T and A244V) in the coagulation
RT   factor VII gene.";
RL   Hum. Mutat. Suppl. 1:S189-S191(1998).
CC   -!- FUNCTION: CIRCULATES IN THE BLOOD IN A ZYMOGEN FORM. FACTOR VII IS
CC       CONVERTED TO FACTOR VIIA BY FACTOR XA, FACTOR XIIA, FACTOR IXA, OR
CC       THROMBIN BY MINOR PROTEOLYSIS. IN THE PRESENCE OF TISSUE FACTOR
CC       AND CALCIUM IONS, FACTOR VIIA THEN CONVERTS FACTOR X TO FACTOR XA
CC       BY LIMITED PROTEOLYSIS. FACTOR VIIA WILL ALSO CONVERT FACTOR IX TO
CC       FACTOR IXA IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM.
CC   -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to
CC       form factor Xa.
CC   -!- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED
CC       BY A DISULFIDE BOND.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC       GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC       CALCIUM.
CC   -!- DISEASE: DEFECTS IN F7 CAN CAUSE COAGULOPATHY.
CC   -!- PHARMACEUTICAL: Available under the names Niastase or Novoseven
CC       (Novo Nordisk). Used for the treatment of bleeding episodes in

  Query Match             50.3%;  Score 99;  DB 1;  Length 466;
  Best Local Similarity   48.8%;  Pred. No. 3.1e-09;
  Matches   20;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
             ||:||  || ||| | |    | |  |: ||:| : |  ||
Db      61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 13
TMG1_HUMAN
ID   TMG1_HUMAN     STANDARD;      PRT;   218 AA.
AC   O14668;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Transmembrane gamma-carboxyglutamic acid protein 1 precursor (Proline-
DE   rich Gla protein 1) (Proline-rich gamma-carboxyglutamic acid protein
DE   1).
GN   PRRG1 OR TMG1 OR PRGP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97404347; PubMed=9256434;
RA   Kulman J.D., Harris J.E., Haldeman B.A., Davie E.W.;
RT   "Primary structure and tissue distribution of two novel proline-rich
RT   gamma-carboxyglutamic acid proteins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:9058-9062(1997).
CC   -!- TISSUE SPECIFICITY: Highly expressed in the spinal cord.
CC   -!- PTM: Gla residues are produced after subsequent posttranslational
CC       modifications of glutamic acid by a vitamin K-dependent gamma-
CC       carboxylase.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF009242; AAB67070.1; -.
DR   HSSP; P00740; 1CFH.
DR   Genew; HGNC:9469; PRRG1.
DR   MIM; 604428; -.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Gamma-carboxyglutamic acid; Vitamin K; Transmembrane.
FT   PROPEP          1     20       POTENTIAL.
FT   CHAIN          21    218       TRANSMEMBRANE GAMMA-CARBOXYGLUTAMIC ACID
FT                                  PROTEIN 1.
FT   DOMAIN         21     83       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       84    106       POTENTIAL.
FT   DOMAIN        107    218       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         24     61       GLA-RICH.
FT   DOMAIN        131    135       POLY-PRO.
SQ   SEQUENCE   218 AA;  24947 MW;  26538A61AB0AEB98 CRC64;

  Query Match             46.7%;  Score 92;  DB 1;  Length 218;
  Best Local Similarity   38.6%;  Pred. No. 2.1e-08;
  Matches   17;  Conservative    7;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           || |   :|||:: | |    | |  |:  ||: : |  ||| :
Db      21 ANGFFEEIRQGNIERECKEEFCTFEEAREAFENNEKTKEFWSTY 64


RESULT 14
THRB_RAT
ID   THRB_RAT       STANDARD;      PRT;   617 AA:
AC   P18292;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=90332426; PubMed=2377469;
RA   Dihanich M., Monard D.;
RT   "cDNA sequence of rat prothrombin.";
RL   Nucleic Acids Res. 18:4251-4251(1990).
RN   [2]
RP   SEQUENCE OF 383-617 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC       fibrinogen to fibrin and releases fibrinopeptide A and B.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC       FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC       BY FACTOR XA.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X52835; CAA37017.1; -.
DR   EMBL; M81397; AAA42240.1; -.
DR   PIR; S10511; S10511.
DR   HSSP; P00734; 1UVS.
DR   MEROPS; S01.217; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL          1     24       POTENTIAL.
FT   PROPEP         25     43
FT   CHAIN          44    617       PROTHROMBIN.
FT   PEPTIDE        44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE       201    323       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN         324    359       THROMBIN LIGHT CHAIN (A).
FT   CHAIN         360    617       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN        109    187       KRINGLE 1.
FT   DOMAIN        215    292       KRINGLE 2.
FT   DOMAIN        360    617       SERINE PROTEASE.
FT   SITE          200    201       CLEAVAGE (BY THROMBIN).
FT   SITE          323    324       CLEAVAGE (BY FACTOR XA).
FT   SITE          359    360       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE      402    402       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE      458    458       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE      564    564       CHARGE RELAY SYSTEM (BY SIMILARITY).
```

```
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   CARBOHYD    120    120       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    412    412       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    552    552       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     61     66       BY SIMILARITY.
FT   DISULFID     91    104       BY SIMILARITY.
FT   DISULFID    109    187       BY SIMILARITY.
FT   DISULFID    130    170       BY SIMILARITY.
FT   DISULFID    158    182       BY SIMILARITY.
FT   DISULFID    215    292       BY SIMILARITY.
FT   DISULFID    236    276       BY SIMILARITY.
FT   DISULFID    264    287       BY SIMILARITY.
FT   DISULFID    332    478       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    387    403       BY SIMILARITY.
FT   DISULFID    532    546       BY SIMILARITY.
FT   DISULFID    560    590       BY SIMILARITY.
SQ   SEQUENCE   617 AA;  70411 MW;  AD27D1B71445DB1D CRC64;

  Query Match           43.9%;  Score 86.5;  DB 1;  Length 617;
  Best Local Similarity  42.2%;  Pred. No. 6.2e-07;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy      1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          ||| ||  ||:|:| | |:   | :  |    |   ||  ||:|:
Db     44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


RESULT 15
THRB_MOUSE
ID   THRB_MOUSE     STANDARD;      PRT;   618 AA.
AC   P19221;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Prothrombin precursor (EC 3.4.21.5).
GN   F2 OR CF2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Liver;
RX   MEDLINE=91025551; PubMed=2222810;
RA   Friezner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA   Fitzgibbon J.J., Pai J.-A., Chapman V.M., Elliott R.W.;
RT   "Characterization of the cDNA coding for mouse prothrombin and
RT   localization of the gene on mouse chromosome 2.";
RL   DNA Cell Biol. 9:487-498(1990).
RN   [2]
RP   SEQUENCE OF 384-618 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92212913; PubMed=1557383;
RA   Banfield D.K., Macgillivray R.T.;
RT   "Partial characterization of vertebrate prothrombin cDNAs:
RT   amplification and sequence analysis of the B chain of thrombin from
RT   nine different species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC   -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC       FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC       AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC       fibrinogen to fibrin and releases fibrinopeptide A and B.
CC   -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC       RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC       ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC       ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC       CHARGED PHOSPHOLIPID SURFACE,WHICH IS ESSENTIAL FOR THE CONVERSION
CC       OF PROTHROMBIN TO THROMBIN.
CC   -!- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC       PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC       FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC       THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC       HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC       V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC       THROMBIN.
CC   -!- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC       FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC       BY FACTOR XA.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X52308; CAA36548.1; -.
DR   EMBL; M81394; AAA40435.1; -.
DR   PIR; A35827; A35827.
DR   HSSP; P00734; 1B7X.
DR   MEROPS; S01.217; -.
DR   MGD; MGI:88380; F2.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003966; Prothrombin.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00051; kringle; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   PRINTS; PR00018; KRINGLE.
DR   PRINTS; PR01505; PROTHROMBIN.
DR   ProDom; PD000395; Kringle; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00130; KR; 2.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 2.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW   Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW   Hydrolase; Serine protease; Kringle; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   PROPEP       25     43
FT   CHAIN        44    618       PROTHROMBIN.
FT   PEPTIDE      44    200       ACTIVATION PEPTIDE (FRAGMENT 1).
FT   PEPTIDE     201    324       ACTIVATION PEPTIDE (FRAGMENT 2).
FT   CHAIN       325    360       THROMBIN LIGHT CHAIN (A).
FT   CHAIN       361    618       THROMBIN HEAVY CHAIN (B).
FT   DOMAIN      109    187       KRINGLE 1.
FT   DOMAIN      215    292       KRINGLE 2.
FT   DOMAIN      361    618       SERINE PROTEASE.
FT   SITE        200    201       CLEAVAGE (BY THROMBIN).
FT   SITE        324    325       CLEAVAGE (BY FACTOR XA).
FT   SITE        360    361       CLEAVAGE (BY FACTOR XA).
FT   ACT_SITE    403    403       CHARGE RELAY SYSTEM (BY SIMILARITY).
```

```
FT   ACT_SITE    459    459       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    565    565       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   MOD_RES      50     50       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      51     51       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      58     58       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      60     60       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      63     63       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      64     64       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      69     69       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      70     70       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      73     73       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES      76     76       GAMMA-CARBOXYGLUTAMIC ACID.
FT   DISULFID     61     66       BY SIMILARITY.
FT   DISULFID     91    104       BY SIMILARITY.
FT   DISULFID    109    187       BY SIMILARITY.
FT   DISULFID    130    170       BY SIMILARITY.
FT   DISULFID    158    182       BY SIMILARITY.
FT   DISULFID    215    293       BY SIMILARITY.
FT   DISULFID    236    276       BY SIMILARITY.
FT   DISULFID    264    288       BY SIMILARITY.
FT   DISULFID    333    479       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    388    404       BY SIMILARITY.
FT   DISULFID    533    547       BY SIMILARITY.
FT   DISULFID    561    591       BY SIMILARITY.
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    413    413       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    553    553       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   618 AA;  70268 MW;  B89F719AAFD601E0 CRC64;

  Query Match            43.9%;  Score 86.5;  DB 1;  Length 618;
  Best Local Similarity  42.2%;  Pred. No. 6.2e-07;
  Matches   19;  Conservative    6;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       1 ANS-FLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||| ||  ||:|:| | |:   | :  |    |   ||  ||:|:
Db      44 ANSGFLEELRKGNLERECVEEQCSYEEAFEALESPQDTDVFWAKY 88


Search completed: May 15, 2003, 13:27:34
Job time : 13 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 15, 2003, 13:24:56 ; Search time 28 Seconds
(without alignments)
323.788 Million cell updates/sec

Title: SEQID1_MOD
Perfect score: 197
Sequence: 1 ANSFLXXLRQGSLXRXCIXX.....XXAKXIFEDVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	151	76.6	456	6	Q9TTR0	Q9ttr0 canis famil
2	140	71.1	460	11	Q91WN8	Q91wn8 mus musculu
3	134	68.0	460	11	Q99PC6	Q99pc6 mus musculu
4	115	58.4	482	11	Q63207	Q63207 rattus norv
5	101	51.3	481	11	O54740	O54740 mus musculu
6	101	51.3	481	11	Q99L32	Q99132 mus musculu
7	101	51.3	481	11	O88947	O88947 mus musculu
8	99	50.3	701	4	Q96PQ8	Q96pq8 homo sapien
9	95	48.2	469	6	Q9GMD9	Q9gmd9 ornithorhyn
10	85	43.1	650	4	Q9NSD0	Q9nsd0 homo sapien
11	85	43.1	650	4	Q16519	Q16519 homo sapien
12	84	42.6	100	4	Q15253	Q15253 homo sapien
13	82.5	41.9	542	5	Q8T6I3	Q8t6i3 halocynthia
14	80	40.6	446	11	Q61109	Q61109 mus musculu
15	80	40.6	456	4	Q14316	Q14316 homo sapien
16	80	40.6	461	6	Q95ND7	Q95nd7 pan troglod
17	80	40.6	461	6	Q95ND6	Q95nd6 pan troglod
18	78	39.6	138	6	Q28994	Q28994 sus scrofa
19	78	39.6	607	13	Q91001	Q91001 gallus gall
20	78	39.6	648	6	Q29094	Q29094 sus scrofa
21	73.5	37.3	433	13	Q90YK1	Q90yk1 brachydanio
22	73	37.1	49	6	Q95ME8	Q95me8 bos taurus
23	73	37.1	399	11	Q9CQW3	Q9cqw3 mus musculu
24	72	36.5	98	13	P82807	P82807 notechis sc
25	72	36.5	608	13	Q9PTW7	Q9ptw7 struthio ca
26	66	33.5	25	11	Q9QVH6	Q9qvh6 rattus sp.
27	65	33.0	179	4	Q8TAS3	Q8tas3 homo sapien
28	65	33.0	198	11	Q8R182	Q8r182 mus musculu
29	65	33.0	673	11	Q61592	Q61592 mus musculu
30	65	33.0	674	11	Q99K57	Q99k57 mus musculu
31	64	32.5	674	11	Q63772	Q63772 rattus sp.
32	63	32.0	678	4	Q14393	Q14393 homo sapien
33	56.5	28.7	459	10	Q9SE22	Q9se22 oryza sativ
34	56.5	28.7	606	10	Q9SJG9	Q9sjg9 arabidopsis
35	56.5	28.7	651	10	Q8S2I8	Q8s2i8 oryza sativ
36	55.5	28.2	575	10	Q94E17	Q94e17 oryza sativ
37	54.5	27.7	567	10	Q8W4J2	Q8w4j2 arabidopsis
38	54.5	27.7	603	10	Q9LPG7	Q9lpg7 arabidopsis
39	53.5	27.2	196	10	O04284	O04284 selaginella
40	53.5	27.2	431	10	Q94EY5	Q94ey5 arabidopsis
41	53.5	27.2	506	10	Q9SPF0	Q9spf0 oryza sativ
42	53.5	27.2	506	10	Q9SE23	Q9se23 oryza sativ
43	53.5	27.2	543	10	Q9MB23	Q9mb23 arabidopsis
44	53.5	27.2	568	10	Q9ASC3	Q9asc3 oryza sativ
45	53.5	27.2	576	10	Q9C9U4	Q9c9u4 arabidopsis

ALIGNMENTS

RESULT 1
Q9TTR0
ID Q9TTR0 PRELIMINARY; PRT; 456 AA.
AC Q9TTR0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein C precursor.
GN PROC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Leeb T., Kopp T., Deppe A., Breen M., Matis U., Brunnberg L.,
RA Brenig B.;
RT "Molecular characterization and chromosomal assignment of the canine
RT protein C gene.";
RL Mamm. Genome 10:135-139(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99371952; PubMed=10443005;
RA Leeb T., Pfeiffer I., Kopp T., Deppe A., Brenig B.;
RT "Analysis of canine protein C gene polymorphisms.";
RL Anim. Genet. 30:237-238(1999).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AJ001979; CAA05126.1; -.
DR HSSP; P04070; 1PCU.
DR MEROPS; S01.218; -.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.

```
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;
KW   Hydroxylation; Repeat; Serine protease; Signal.
FT   SIGNAL        1     42       POTENTIAL.
FT   CHAIN        43    192       PROTEIN C LIGHT CHAIN.
FT   CHAIN       193    194       PROTEIN C CONNECTING DIPEPTIDE.
FT   CHAIN       195    456       PROTEIN C HEAVY CHAIN.
SQ   SEQUENCE   456 AA;  50813 MW;  7AD3A8C1C34E59FF CRC64;

  Query Match             76.6%;  Score 151;  DB 6;  Length 456;
  Best Local Similarity   63.6%;  Pred. No. 1.6e-17;
  Matches   28;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  :| ||| | |:  ||||  || ||::|||||:|||:
Db      43 ANSFLEEIRAGSLERECMEEICDFEEAKEIFQNVDDTLAYWSKY 86


RESULT 2
Q91WN8
ID   Q91WN8      PRELIMINARY;      PRT;   460 AA.
AC   Q91WN8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to protein C.
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013896; AAH13896.1; -.
DR   MGD; MGI:97771; Proc.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_2.
DR   PROSITE; PS01187; EGF_CA; UNKNOWN_1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   460 AA;  51818 MW;  0117F26E68FCC274 CRC64;

  Query Match             71.1%;  Score 140;  DB 11;  Length 460;
  Best Local Similarity   59.1%;  Pred. No. 1.3e-15;
  Matches   26;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  :| ||| | |:  ||||  |: ||::|:|||||| |:
Db      42 ANSFLEEMRPGSLERECMEEICDFEEAQEIFQNVEDTLAFWIKY 85


RESULT 3
Q99PC6
ID   Q99PC6      PRELIMINARY;      PRT;   460 AA.
AC   Q99PC6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Anticoagulant protein C.
GN   PROC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL;
RA   Korf I.;
RT   "Complete sequence of UC72A01.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AF318182; AAK07918.1; -.
DR   HSSP; P04070; 1PCU.
DR   MEROPS; S01.218; -.
DR   MGD; MGI:97771; Proc.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00001; EGF_like; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;
KW   Hydroxylation; Repeat; Serine protease.
SQ   SEQUENCE   460 AA;  51784 MW;  0293BC25E9D3ED16 CRC64;

  Query Match             68.0%;  Score 134;  DB 11;  Length 460;
  Best Local Similarity   56.8%;  Pred. No. 1.4e-14;
  Matches   25;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  :| ||| | |:  |||   |: ||::|:|||||| |:
Db      42 ANSFLEEMRPGSLERECMEEICDLEEAQEIFQNVEDTLAFWIKY 85


RESULT 4
Q63207
ID   Q63207      PRELIMINARY;      PRT;   482 AA.
AC   Q63207;
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Factor X.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=96093366; PubMed=8578539;
RA   Stanton C., Ross R.P., Hutson S., Wallin R.;
RT   "Evidence for competition between vitamin K-dependent clotting factors
RT   for intracellular processing by the vitamin K-dependent gamma-
RT   carboxylase.";
RL   Thromb. Res. 80:63-73(1995).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; X79807; CAA56202.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;
KW   Hydroxylation; Repeat; Serine protease.
SQ   SEQUENCE   482 AA;  54265 MW;  0284678E3954A698 CRC64;

  Query Match             58.4%;  Score 115;  DB 11;  Length 482;
  Best Local Similarity   43.2%;  Pred. No. 2.8e-11;
  Matches   19;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||||   :::|:| | |:  || |  |: :||| : |  ||:|:
Db      41 ANSFFEEIKKGNLERECVEEICSFEEAREVFEDNEKTTEFWNKY 84


RESULT 5
O54740
ID   O54740      PRELIMINARY;      PRT;   481 AA.
AC   O54740;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Coagulation factor X precursor (EC 3.4.21.6).
GN   F10 OR FA10.
OS   Mus musculus (Mouse).
OG   Plasmid pBluescript.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=98454993; PubMed=9783672;
RA   Heidtmann H.H., Kontermann R.E.;
RT   "Cloning and recombinant expression of mouse coagulation factor X.";
RL   Thromb. Res. 92:33-41(1998).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AJ222677; CAA10933.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Plasmid;
KW   Repeat; Serine protease; Signal.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41    481       COAGULATION FACTOR X.
SQ   SEQUENCE   481 AA;  53986 MW;  CF702DE5EF9D97AE CRC64;

  Query Match             51.3%;  Score 101;  DB 11;  Length 481;
  Best Local Similarity   38.6%;  Pred. No. 7.4e-09;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||||   :::|:| | |:  || |  |: :||| : |  ||:|:
Db      41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


RESULT 6
Q99L32
ID   Q99L32      PRELIMINARY;      PRT;   481 AA.
AC   Q99L32;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Coagulation factor X.
GN   F10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
```

```
DR   EMBL; BC003877; AAH03877.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00001; EGF_like; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW   Serine protease.
SQ   SEQUENCE   481 AA;  54004 MW;  BD88E96C8A0B7E7F CRC64;

  Query Match             51.3%;  Score 101;  DB 11;  Length 481;
  Best Local Similarity   38.6%;  Pred. No. 7.4e-09;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||||    ::|:| | |:  || :   : |||| : |  :|:|:
Db      41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


RESULT 7
O88947
ID   O88947      PRELIMINARY;      PRT;   481 AA.
AC   O88947;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Coagulation factor X precursor.
GN   F10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL6 X CBA; TISSUE=LIVER;
RX   MEDLINE=98347933; PubMed=9684791;
RA   Liang Z., Cooper A., DeFord M.E., Carmeliet P., Collen D.,
RA   Castellino F.J., Rosen E.D.;
RT   "Cloning and characterization of a cDNA encoding murine coagulation
RT   factor X.";
RL   Thromb. Haemost. 80:87-91(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SJ;
RA   Cooper A., Liang Z., Castellino F.J., Rosen E.D.;
RT   "Cloning and Characterization of the Murine Factor X Gene.";
RL   Thromb. Haemost. 0:0-0(2000).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AF087644; AAC36345.1; -.
DR   EMBL; AF211347; AAF22980.1; -.
DR   HSSP; P00742; 1XKA.
DR   MEROPS; S01.216; -.
DR   MGD; MGI:103107; F10.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW   Serine protease; Signal.
FT   SIGNAL        1     40       POTENTIAL.
FT   CHAIN        41    481       COAGULATION FACTOR X.
SQ   SEQUENCE   481 AA;  54018 MW;  8AC09DE5EF9D271E CRC64;

  Query Match             51.3%;  Score 101;  DB 11;  Length 481;
  Best Local Similarity   38.6%;  Pred. No. 7.4e-09;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||||    ::|:| | |:  || :   : |||| : |  :|:|:
Db      41 ANSFFEEFKKGNLERECMEEICSYEEVREIFEDDEKTKEYWTKY 84


RESULT 8
Q96PQ8
ID   Q96PQ8      PRELIMINARY;      PRT;   701 AA.
AC   Q96PQ8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Factor VII active site mutant immunoconjugate.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21477448; PubMed=11593034;
RA   Hu Z., Garen A.;
RT   "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT   cells for immunotherapy in mouse models of prostatic cancer.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR   EMBL; AF272774; AAK58686.1; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
```

```
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00181; EGF; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PROSITE; PS01187; EGF_CA; UNKNOWN_1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   701 AA;  77826 MW;  94AC6CEB42CC992F CRC64;

  Query Match             50.3%;  Score 99;  DB 4;  Length 701;
  Best Local Similarity   48.8%;  Pred. No. 2.5e-08;
  Matches   20;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
           ||:||  || ||| | |     | |  |: ||:| : |  ||
Db      61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 9
Q9GMD9
ID   Q9GMD9      PRELIMINARY;      PRT;   469 AA.
AC   Q9GMD9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Coagulation factor X.
OS   Ornithorhynchus anatinus (Duckbill platypus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX   NCBI_TaxID=9258;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21015017; PubMed=11132153;
RA   Poorafshar M., Aveskogh M., Munday B., Hellman L.;
RT   "Identification and structural analysis of four serine proteases in a
RT   monotreme, the platypus, Ornithorhynchus anatinus.";
RL   Immunogenetics 52:19-28(2000).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AF275654; AAG00453.1; -.
DR   HSSP; P00742; 1XKB.
DR   MEROPS; S01.216; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 2.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   469 AA;  52196 MW;  4C66C230D0758F6A CRC64;

  Query Match             48.2%;  Score 95;  DB 6;  Length 469;
  Best Local Similarity   40.5%;  Pred. No. 7.9e-08;
  Matches   17;  Conservative    7;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWS 42
           |||    |::|:| | |    | :  |: :||| | |  ||:
Db      41 ANSLFEELKKGNLERECNEETCSYEEAREVFEDTDKTNEFWN 82


RESULT 10
Q9NSD0
ID   Q9NSD0      PRELIMINARY;      PRT;   650 AA.
AC   Q9NSD0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Protein S precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Wydro R., Cohen E., Dackowski W., Stenflo J., Lundwall A.,
RA   Dahlback B.;
RL   Submitted (FEB-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X12892; CAA31383.1; -.
DR   HSSP; P00740; 1CFH.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 4.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00054; laminin_G; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 3.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00282; LamG; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS01187; EGF_CA; 2.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat;
KW   Signal.
FT   SIGNAL        1     15       POTENTIAL.
FT   CHAIN        16    650       POTENTIAL.
SQ   SEQUENCE   650 AA;  72480 MW;  C67345ECE8645174 CRC64;

  Query Match             43.1%;  Score 85;  DB 4;  Length 650;
  Best Local Similarity   38.6%;  Pred. No. 6.1e-06;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||| |   :||:| | ||  :|:   |: :||:  :|  |: |:
Db      16 ANSLLEETKQGNLERECIEELCNKEEAREVFENDPETDYFYPKY 59


RESULT 11
Q16519
ID   Q16519      PRELIMINARY;      PRT;   650 AA.
AC   Q16519;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Protein S precursor (Fragment).
GN   PROS1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86313649; PubMed=2944113;
RA   Lundwall A., Dackowski W., Cohen E., Shaffer M., Mahr A., Dahlback B.,
RA   Stenflo J., Wydro R.;
RT   "Isolation and sequence of the cDNA for human protein S, a regulator
RT   of blood coagulation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:6716-6720(1986).
DR   EMBL; M14338; AAA60181.1; -.
DR   HSSP; P00740; 1CFH.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 4.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00054; laminin_G; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 3.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00282; LamG; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS01187; EGF_CA; 2.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat;
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     15       POTENTIAL.
FT   CHAIN        16    650       PROTEIN S.
SQ   SEQUENCE   650 AA;  72462 MW;  9A8C044C503BF474 CRC64;

  Query Match             43.1%;  Score 85;  DB 4;  Length 650;
  Best Local Similarity   38.6%;  Pred. No. 6.1e-06;
  Matches   17;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||  |   :||:| | ||  :|:   |: :||:  :|  |: |:
Db     16 ANSLLEETKQGNLERECIEELCNKEEAREVFENDPETDYFYPKY 59


RESULT 12
Q15253
ID   Q15253      PRELIMINARY;      PRT;   100 AA.
AC   Q15253;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Thrombin precursor (Fragment).
GN   F2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87182874; PubMed=3471151;
RA   MacGillivray R.T., Irwin D.M., Guinto E.R., Stone J.C.;
RT   "Recombinant genetic approaches to functional mapping of thrombin.";
RL   Ann. N. Y. Acad. Sci. 485:73-79(1986).
DR   EMBL; M33031; AAA60220.1; -.
DR   HSSP; P00735; 2PF1.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00594; gla; 1.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00069; GLA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
KW   Signal.
FT   SIGNAL        1     43       POTENTIAL.
FT   CHAIN        44   >100       POTENTIAL.
FT   NON_TER     100    100
SQ   SEQUENCE   100 AA;  11302 MW;  FD0E5D0174E1F6FE CRC64;

  Query Match             42.6%;  Score 84;  DB 4;  Length 100;
  Best Local Similarity   36.4%;  Pred. No. 1.2e-06;
  Matches   16;  Conservative    8;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          ||:||  :|:|:| | |:   | :  |    |    |  ||:|:
Db     44 ANTFLEEVRKGNLERECVEETCSYEEAFEALESSTATDVFWAKY 87


RESULT 13
Q8T6I3
ID   Q8T6I3      PRELIMINARY;      PRT;   542 AA.
AC   Q8T6I3;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Gla-like protein.
OS   Halocynthia roretzi (Sea squirt).
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC   Stolidobranchia; Pyuridae; Halocynthia.
OX   NCBI_TaxID=7729;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang C.-P., Stafford D.W.;
RT   "Halocynthia roretzi gla-like protein partial genomic DNA sequence.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF466701; AAL74247.2; -.
SQ   SEQUENCE   542 AA;  62090 MW;  EB9BF13FE42B32FE CRC64;

  Query Match             41.9%;  Score 82.5;  DB 5;  Length 542;
  Best Local Similarity   34.9%;  Pred. No. 1.3e-05;
  Matches   15;  Conservative   10;  Mismatches   17;  Indels    1;  Gaps    1;

Qy      3 SFLXXLRQGSLXRXCIXXICDFXXAKXIFE-DVDDTLAFWSKH 44
          |    ::||:| | |   :| |  |: :|| :: |   ||:|:
Db     33 SHFEEIQQGNLERECYEELCSFEEAREVFETNIQDLNEFWAKY 75


RESULT 14
Q61109
ID   Q61109      PRELIMINARY;      PRT;   446 AA.
AC   Q61109;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Coagulation factor VII.
GN   F7 OR FVII.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=96276538; PubMed=8701412;
RA   Idusogie E., Rosen E., Geng J.P., Carmeliet P., Collen D.,
RA   Castellino F.J.;
RT   "Characterization of a cDNA encoding murine coagulation factor VII.";
RL   Thromb. Haemost. 75:481-487(1996).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
```

```
DR   EMBL; U44795; AAC52570.1; -.
DR   HSSP; P08709; 1FAK.
DR   MEROPS; S01.215; -.
DR   MGD; MGI:109325; F7.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW   Serine protease.
SQ   SEQUENCE   446 AA;  50318 MW;  482FD09BEFDA6870 CRC64;

  Query Match             40.6%;  Score 80;  DB 11;  Length 446;
  Best Local Similarity   43.9%;  Pred. No. 2.9e-05;
  Matches   18;  Conservative    3;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
           |||  |   |  |||  | |    | |  |: ||:  : |  ||
Db      42 ANSLLEELWPGSLERECNEEQCSFEEAREIFKSPERTKQFW 82


RESULT 15
Q14316
ID   Q14316      PRELIMINARY;      PRT;   456 AA.
AC   Q14316;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   F9 (Coagulation factor IX (Plasma THROMBOPLASTIC component, christmas
DE   disease,HAEMOPHILIA B)) (Factor IX).
GN   F9 OR FACTOR IX.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bird C.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 3-19 FROM N.A.
RX   MEDLINE=88327116; PubMed=3416069;
RA   Reitsma P.A., Bertina R.M., Ploos van Amstel J.K., Riemens A.,
RA   Briet E.;
RT   "The putative factor IX gene promoter in hemophilia B Leyden.";
RL   Blood 72:1074-1076(1988).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   EMBL; AL033403; CAA21954.1; -.
DR   EMBL; X55008; CAB38245.2; -.
DR   HSSP; P00740; 1CFH.
DR   MEROPS; S01.214; -.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00594; gla; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW   Serine protease.
SQ   SEQUENCE   456 AA;  51149 MW;  54E20A1B3964E234 CRC64;

  Query Match             40.6%;  Score 80;  DB 4;  Length 456;
  Best Local Similarity   37.1%;  Pred. No. 3e-05;
  Matches   13;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      10 QGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           ||:| | |:   | |  |: :||: : |  || ::
Db      52 QGNLERECMEEKCSFEEAREVFENTERTTEFWKQY 86


Search completed: May 15, 2003, 13:28:11
Job time : 31 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 15, 2003, 13:26:41 ; Search time 15 Seconds
(without alignments)
86.307 Million cell updates/sec

Title: SEQID1_MOD
Perfect score: 197
Sequence: 1 ANSFLXXLRQGSLXRXCIXX.....XXAKXIFEDVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	179	90.9	44	3	US-08-955-636-24	Sequence 24, Appl
2	176	89.3	44	3	US-08-955-636-35	Sequence 35, Appl
3	173	87.8	44	3	US-08-955-636-20	Sequence 20, Appl
4	170	86.3	44	3	US-08-955-636-21	Sequence 21, Appl
5	168	85.3	44	3	US-08-955-636-19	Sequence 19, Appl
6	168	85.3	44	3	US-08-955-636-22	Sequence 22, Appl
7	160	81.2	44	3	US-08-955-636-1	Sequence 1, Appli
8	160	81.2	44	3	US-08-955-636-25	Sequence 25, Appl
9	160	81.2	45	2	US-08-965-832-2	Sequence 2, Appli
10	160	81.2	419	1	US-08-295-411-1	Sequence 1, Appli
11	160	81.2	419	2	US-08-955-471-1	Sequence 1, Appli
12	160	81.2	419	4	US-09-667-570A-3	Sequence 3, Appli
13	160	81.2	419	5	PCT-US92-10242-1	Sequence 1, Appli
14	160	81.2	460	2	US-08-756-506-2	Sequence 2, Appli
15	160	81.2	460	2	US-08-756-506-4	Sequence 4, Appli
16	160	81.2	460	6	5270178-13	Patent No. 5270178
17	160	81.2	460	6	5270178-14	Patent No. 5270178
18	160	81.2	460	6	5270178-15	Patent No. 5270178
19	160	81.2	460	6	5270178-16	Patent No. 5270178
20	160	81.2	461	6	5225537-2	Patent No. 5225537
21	160	81.2	461	6	5270178-17	Patent No. 5270178
22	160	81.2	461	6	5270178-18	Patent No. 5270178
23	160	81.2	461	6	5460953-3	Patent No. 5460953
24	147	74.6	42	2	US-08-745-254A-2	Sequence 2, Appli
25	147	74.6	461	6	5270178-2	Patent No. 5270178
26	143	72.6	41	1	US-08-229-280-5	Sequence 5, Appli
27	129	65.5	409	4	US-09-065-872-2	Sequence 2, Appli
28	129	65.5	409	4	US-09-667-570A-2	Sequence 2, Appli
29	129	65.5	410	4	US-09-065-872-1	Sequence 1, Appli
30	129	65.5	410	4	US-09-667-570A-1	Sequence 1, Appli
31	117	59.4	44	3	US-08-955-636-23	Sequence 23, Appl
32	116	58.9	44	3	US-08-955-636-2	Sequence 2, Appli
33	114	57.9	139	1	US-08-330-978-2	Sequence 2, Appli
34	114	57.9	139	1	US-08-474-042-2	Sequence 2, Appli
35	114	57.9	139	1	US-08-484-558-2	Sequence 2, Appli
36	114	57.9	139	1	US-08-774-592-2	Sequence 2, Appli
37	114	57.9	437	1	US-08-487-037-2	Sequence 2, Appli
38	114	57.9	437	1	US-08-487-037-3	Sequence 3, Appli
39	114	57.9	487	1	US-08-469-486-53	Sequence 53, Appl
40	114	57.9	487	2	US-08-469-658-53	Sequence 53, Appl
41	114	57.9	488	1	US-08-487-037-1	Sequence 1, Appli
42	114	57.9	492	1	US-08-469-486-2	Sequence 2, Appli
43	114	57.9	492	2	US-08-469-658-2	Sequence 2, Appli
44	110	55.8	448	1	US-08-295-411-3	Sequence 3, Appli
45	110	55.8	448	2	US-08-955-471-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-955-636-24
; Sequence 24, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary
; TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 09531/002001
; CURRENT APPLICATION NUMBER: US/08/955,636A
; CURRENT FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-24

Query Match 90.9%; Score 179; DB 3; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
||||||||||||||||||||||||||||||||||||||||||||
Db 1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44

RESULT 2
US-08-955-636-35
; Sequence 35, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
; APPLICANT: Nelsestuen, Gary
; TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 09531/002001
; CURRENT APPLICATION NUMBER: US/08/955,636A
; CURRENT FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-35

  Query Match             89.3%;  Score 176;  DB 3;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 3.8e-23;
  Matches   43;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||:|||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLREGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 3
US-08-955-636-20
; Sequence 20, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
  APPLICANT: Nelsestuen, Gary
  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
  TITLE OF INVENTION: POLYPEPTIDES
  FILE REFERENCE: 09531/002001
  CURRENT APPLICATION NUMBER: US/08/955,636A
  CURRENT FILING DATE: 1997-10-23
  NUMBER OF SEQ ID NOS: 35
  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
  LENGTH: 44
  TYPE: PRT
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: MOD_RES
  LOCATION: (0)...(0)
  OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-20

  Query Match             87.8%;  Score 173;  DB 3;  Length 44;
  Best Local Similarity   97.7%;  Pred. No. 1.2e-22;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||| ||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRQSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 4
US-08-955-636-21
; Sequence 21, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-21

  Query Match             86.3%;  Score 170;  DB 3;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 4e-22;
  Matches   42;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||: |||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRESSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 5
US-08-955-636-19
; Sequence 19, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-19

  Query Match             85.3%;  Score 168;  DB 3;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 8.9e-22;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  ||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 6
US-08-955-636-22
; Sequence 22, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-22

  Query Match             85.3%;  Score 168;  DB 3;  Length 44;
  Best Local Similarity   95.5%;  Pred. No. 8.9e-22;
  Matches   42;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  ||||||||||||||||||||||||||||||||||
Db      1 ANSFLXXLRDSSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44


RESULT 7
US-08-955-636-1
; Sequence 1, Application US/08955636A
; Patent No. 6017882
```

```
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-1

  Query Match             81.2%;  Score 160;  DB 3;  Length 44;
  Best Local Similarity   90.9%;  Pred. No. 2e-20;
  Matches   40;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps .  0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  ||||||||||||||||||::||||||||||
Db      1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 8
US-08-955-636-25
; Sequence 25, Application US/08955636A
; Patent No. 6017882
; GENERAL INFORMATION:
;  APPLICANT: Nelsestuen, Gary
;  TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 09531/002001
;  CURRENT APPLICATION NUMBER: US/08/955,636A
;  CURRENT FILING DATE: 1997-10-23
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
;   LENGTH: 44
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Xaa=gamma carboxyglutamic acid or glutamic acid
US-08-955-636-25

  Query Match             81.2%;  Score 160;  DB 3;  Length 44;
  Best Local Similarity   93.2%;  Pred. No. 2e-20;
  Matches   41;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  |||||||||||||||| ||||||||||||||||
Db      1 ANSFLXXLRHSSLXRXCIXXICDFXXAFXIFEDVDDTLAFWSKH 44


RESULT 9
US-08-965-832-2
; Sequence 2, Application US/08965832
; Patent No. 5847085
;  GENERAL INFORMATION:
;    APPLICANT:  CHARLES T. ESMON AND MIKHAIL D. SMIRNOV
;    TITLE OF INVENTION:  Modified Protein C
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center, 1201 West
;      STREET:  Peachtree Street
;      CITY:  Atlanta
;      STATE:  GA
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/965,832
;      FILING DATE:  7-NOV-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/745,254
;      FILING DATE:  8-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/053,768
;      FILING DATE:  25-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF 165/167
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-873-8794
;      TELEFAX:  (404)-873-8795
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:
;      LOCATION:  6, 7, 14, 16, 19, 20, 25, 26, 29
;      OTHER INFORMATION:  /note= "where Xaa means gamma
;      OTHER INFORMATION:  carboxylglutamic acid"
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      OTHER INFORMATION:  /note= "partial sequence of human protein C"
US-08-965-832-2

  Query Match             81.2%;  Score 160;  DB 2;  Length 45;
  Best Local Similarity   90.9%;  Pred. No. 2.1e-20;
  Matches   40;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||||||  |||||||||||||||||||::||||||||||
Db      1 ANSFLXXLRHSSLXRXCIXXICDFXXAKXIFQNVDDTLAFWSKH 44


RESULT 10
US-08-295-411-1
; Sequence 1, Application US/08295411
; Patent No. 5679639
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5679639th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/295,411
;      FILING DATE:  22-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/793,989
;      FILING DATE:  18-NOV-1991
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI263.0C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  419 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..157
;      OTHER INFORMATION:  /note= "Protein C Light Chain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  158..169
;      OTHER INFORMATION:  /note= "Protein C Activation
;      OTHER INFORMATION:  Peptide"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  170..419
;      OTHER INFORMATION:  /note= "Protein C Heavy Chain"
US-08-295-411-1

  Query Match             81.2%;  Score 160;  DB 1;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 2.7e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  ||  || | ||  ||||  || ||::||||||||||||
Db       1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 11
US-08-955-471-1
; Sequence 1, Application US/08955471
; Patent No. 5968751
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5968751th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,471
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/295,411
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI263.0C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  419 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..157
;      OTHER INFORMATION:  /note= "Protein C Light Chain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  158..169
;      OTHER INFORMATION:  /note= "Protein C Activation
;      OTHER INFORMATION:  Peptide"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  170..419
;      OTHER INFORMATION:  /note= "Protein C Heavy Chain"
US-08-955-471-1

  Query Match             81.2%;  Score 160;  DB 2;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 2.7e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  ||  || | ||  ||||  || ||::||||||||||||
Db       1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 12
US-09-667-570A-3
; Sequence 3, Application US/09667570A
; Patent No. 6436397
; GENERAL INFORMATION:
;  APPLICANT: Baker, Jeffrey C
;  APPLICANT:  Carlson, Andrew D
;  APPLICANT:  Huang, Lihua
;  APPLICANT:  Sheliga, Theodore  A
;  TITLE OF INVENTION: Improved Methods for Processing Activated Protein C
;  FILE REFERENCE: X-11796A
;  CURRENT APPLICATION NUMBER: US/09/667,570A
;  CURRENT FILING DATE:  2000-09-21
;  PRIOR APPLICATION NUMBER: 60/045,255
;  PRIOR FILING DATE: 1997-04-28
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-667-570A-3

  Query Match             81.2%;  Score 160;  DB 4;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 2.7e-19;
```

Matches 31; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

```
Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          ||||| || || | ||  |||| || ||::||||||||||||
Db      1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44
```

RESULT 13
PCT-US92-10242-1
; Sequence 1, Application PC/TUS9210242
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Mesters, Rolf
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10242
; FILING DATE: 19921118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,989
; FILING DATE: 18-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR0472P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..157
; OTHER INFORMATION: /note= "Protein C Light Chain"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 158..169
; OTHER INFORMATION: /note= "Protein C Activation
; OTHER INFORMATION: Peptide"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 170..419
; OTHER INFORMATION: /note= "Protein C Heavy Chain"
PCT-US92-10242-1

Query Match 81.2%; Score 160; DB 5; Length 419;
Best Local Similarity 70.5%; Pred. No. 2.7e-19;
Matches 31; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

```
Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          ||||| || || | ||  |||| || ||::||||||||||||
Db      1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44
```

RESULT 14
US-08-756-506-2
; Sequence 2, Application US/08756506
; Patent No. 5905185
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Temperley, Simon M.
; APPLICANT: Foster, Donald C.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Prunkard, Donna E.
; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
; TITLE OF INVENTION: ANIMALS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,506
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-756-506-2

Query Match 81.2%; Score 160; DB 2; Length 460;
Best Local Similarity 70.5%; Pred. No. 3.1e-19;
Matches 31; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

```
Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          ||||| || || | ||  |||| || ||::||||||||||||
Db     43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86
```

RESULT 15
US-08-756-506-4
; Sequence 4, Application US/08756506
; Patent No. 5905185
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Temperley, Simon M.
; APPLICANT: Foster, Donald C.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Prunkard, Donna E.
; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
; TITLE OF INVENTION: ANIMALS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.

```
;       STREET:  1201 Eastlake Avenue East
;       CITY:  Seattle
;       STATE:  WA
;       COUNTRY:  USA
;       ZIP:  98102
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/756,506
;       FILING DATE:
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sawislak, Deborah A
;       REGISTRATION NUMBER:  37,438
;       REFERENCE/DOCKET NUMBER:  95-28
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  206-442-6672
;       TELEFAX:  206-442-6678
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  460 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-756-506-4

  Query Match             81.2%;  Score 160;  DB 2;  Length 460;
  Best Local Similarity   70.5%;  Pred. No. 3.1e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           |||||  ||  || | ||  ||||  || ||::||||||||||||
Db      43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


Search completed: May 15, 2003, 13:28:55
Job time : 16 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 15, 2003, 13:27:21 ; Search time 19 Seconds
(without alignments)
223.360 Million cell updates/sec

Title: SEQID1_MOD
Perfect score: 197
Sequence: 1 ANSFLXXLRQGSLXRXCIXX.....XXAKXIFEDVDDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	179	90.9	419	9	US-10-182-263-6	Sequence 6, Appli
2	174	88.3	419	9	US-10-182-263-3	Sequence 3, Appli
3	174	88.3	419	9	US-10-182-263-4	Sequence 4, Appli
4	174	88.3	419	9	US-10-182-263-5	Sequence 5, Appli
5	160	81.2	419	9	US-10-182-263-1	Sequence 1, Appli
6	160	81.2	419	9	US-09-978-917A-4	Sequence 4, Appli
7	160	81.2	461	9	US-10-182-263-2	Sequence 2, Appli
8	160	81.2	461	9	US-09-978-917A-2	Sequence 2, Appli
9	99	50.3	466	9	US-10-017-122-2	Sequence 2, Appli
10	96	48.7	406	9	US-10-109-498-1	Sequence 1, Appli
11	84.5	42.9	96	9	US-09-759-130B-313	Sequence 313, App
12	84.5	42.9	96	9	US-10-189-123-43	Sequence 43, Appl
13	84.5	42.9	209	9	US-09-759-130B-312	Sequence 312, App
14	84.5	42.9	209	9	US-10-189-123-42	Sequence 42, Appl
15	84.5	42.9	226	9	US-09-759-130B-310	Sequence 310, App
16	84.5	42.9	226	9	US-10-189-123-40	Sequence 40, Appl
17	80	40.6	415	10	US-09-118-748-2	Sequence 2, Appli
18	80	40.6	461	9	US-10-132-829-5	Sequence 5, Appli
19	80	40.6	461	10	US-09-884-901-3	Sequence 3, Appli
20	64.5	32.7	95	9	US-09-759-130B-356	Sequence 356, App
21	64.5	32.7	95	9	US-10-189-123-86	Sequence 86, Appl
22	64.5	32.7	208	9	US-09-759-130B-355	Sequence 355, App
23	64.5	32.7	208	9	US-10-189-123-85	Sequence 85, Appl
24	64.5	32.7	225	9	US-09-759-130B-353	Sequence 353, App
25	64.5	32.7	225	9	US-10-189-123-83	Sequence 83, Appl
26	49	24.9	1363	9	US-09-375-248-19	Sequence 19, Appl
27	48	24.4	348	10	US-09-982-610-18	Sequence 18, Appl
28	48	24.4	1298	10	US-09-982-610-33	Sequence 33, Appl
29	48	24.4	1363	9	US-09-375-248-2	Sequence 2, Appli
30	47	23.9	180	10	US-09-766-678-6	Sequence 6, Appli
31	47	23.9	317	9	US-09-939-833-5	Sequence 5, Appli
32	47	23.9	317	10	US-09-939-754-5	Sequence 5, Appli
33	47	23.9	317	10	US-09-939-832-5	Sequence 5, Appli
34	47	23.9	367	9	US-09-939-833-12	Sequence 12, Appl
35	47	23.9	367	10	US-09-939-754-12	Sequence 12, Appl
36	47	23.9	367	10	US-09-939-832-12	Sequence 12, Appl
37	47	23.9	1356	9	US-09-969-037-7	Sequence 7, Appli
38	47	23.9	1356	9	US-10-022-939-2	Sequence 2, Appli
39	47	23.9	1356	9	US-10-100-405A-2	Sequence 2, Appli
40	47	23.9	1367	10	US-09-919-408-6	Sequence 6, Appli
41	47	23.9	1367	10	US-09-766-678-2	Sequence 2, Appli
42	47	23.9	1367	10	US-09-872-136-6	Sequence 6, Appli
43	45	22.8	254	10	US-09-796-149-4	Sequence 4, Appli
44	44.5	22.6	49	9	US-09-836-392-34	Sequence 34, Appl
45	44	22.3	52	9	US-09-796-692-2179	Sequence 2179, Ap

ALIGNMENTS

RESULT 1
US-10-182-263-6
; Sequence 6, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-6

Query Match 90.9%; Score 179; DB 9; Length 419;
Best Local Similarity 79.5%; Pred. No. 5.4e-22;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  ||||||  |  ||  ||||  ||  ||||||||||||||
Db      1 ANSFLEELRQGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44
```

RESULT 2
US-10-182-263-3
; Sequence 3, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES

```
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-3

  Query Match             88.3%;  Score 174;  DB 9;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 3.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  || ||| | ||  |||| || ||||||||||||||||
Db      1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 3
US-10-182-263-4
; Sequence 4, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-4

  Query Match             88.3%;  Score 174;  DB 9;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 3.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  || ||| | ||  |||| || ||||||||||||||||
Db      1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 4
US-10-182-263-5
; Sequence 5, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-5

  Query Match             88.3%;  Score 174;  DB 9;  Length 419;
  Best Local Similarity   77.3%;  Pred. No. 3.8e-21;
  Matches   34;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  || ||| | ||  |||| || ||||||||||||||||
Db      1 ANSFLEELRHGSLERECIEEICDFEEAKEIFEDVDDTLAFWSKH 44


RESULT 5
US-10-182-263-1
; Sequence 1, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-1

  Query Match             81.2%;  Score 160;  DB 9;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 8.6e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  ||  || | ||  |||| || ||::||||||||||||
Db      1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 6
US-09-978-917A-4
; Sequence 4, Application US/09978917A
; Publication No. US20030027299A1
; GENERAL INFORMATION:
;  APPLICANT: Maxygen Aps; Maxygen Holdings
;  TITLE OF INVENTION: Protein C or activated protein C-like molecules
;  FILE REFERENCE: 0219us310 - protein C
;  CURRENT APPLICATION NUMBER: US/09/978,917A
;  CURRENT FILING DATE:  2001-10-17
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-978-917A-4

  Query Match             81.2%;  Score 160;  DB 9;  Length 419;
  Best Local Similarity   70.5%;  Pred. No. 8.6e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
```

```
          ||||| ||  || | ||  |||| || ||::|||||||||||
Db      1 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 44


RESULT 7
US-10-182-263-2
; Sequence 2, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
;  APPLICANT: Gerlitz, Bruce E
;  APPLICANT:  Jones, Bryan E
;  APPLICANT:  Grinnell, Brian W
;  TITLE OF INVENTION: PROTEIN C DERIVATIVES
;  FILE REFERENCE: X-13611
;  CURRENT APPLICATION NUMBER: US/10/182,263
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: 60/181948
;  PRIOR FILING DATE: 2002-02-11
;  PRIOR APPLICATION NUMBER: 60/189199
;  PRIOR FILING DATE: 2000-03-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-182-263-2

  Query Match             81.2%;  Score 160;  DB 9;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 9.5e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  ||  || | ||  |||| || ||::|||||||||||
Db     43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 8
US-09-978-917A-2
; Sequence 2, Application US/09978917A
; Publication No. US20030027299A1
; GENERAL INFORMATION:
;  APPLICANT: Maxygen Aps; Maxygen Holdings
;  TITLE OF INVENTION: Protein C or activated protein C-like molecules
;  FILE REFERENCE: 0219us310 - protein C
;  CURRENT APPLICATION NUMBER: US/09/978,917A
;  CURRENT FILING DATE:  2001-10-17
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: (1)...(42)
;   FEATURE:
;   NAME/KEY: CHAIN
;   LOCATION: (43)...(461)
US-09-978-917A-2

  Query Match             81.2%;  Score 160;  DB 9;  Length 461;
  Best Local Similarity   70.5%;  Pred. No. 9.5e-19;
  Matches   31;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
          |||||  ||  || | ||  |||| || ||::|||||||||||
Db     43 ANSFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKH 86


RESULT 9
US-10-017-122-2
; Sequence 2, Application US/10017122
; Publication No. US20030087244A1
; GENERAL INFORMATION:
;  APPLICANT: McCarthy, Jeanette
;  TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
;  FILE REFERENCE: MMI-007
;  CURRENT APPLICATION NUMBER: US/10/017,122
;  CURRENT FILING DATE:  2001-12-14
;  PRIOR APPLICATION NUMBER: 60/327,487
;  PRIOR FILING DATE: 2001-10-09
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 466
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-017-122-2

  Query Match             50.3%;  Score 99;  DB 9;  Length 466;
  Best Local Similarity   48.8%;  Pred. No. 1.8e-08;
  Matches   20;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
          ||:||  || ||| | |    | |  |: ||:| : |  ||
Db     61 ANAFLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFW 101


RESULT 10
US-10-109-498-1
; Sequence 1, Application US/10109498
; Publication No. US20030044908A1
; GENERAL INFORMATION:
;  APPLICANT: Persson, Egon
;  TITLE OF INVENTION: Coagulation Factor VII Derivatives
;  FILE REFERENCE: 6286.200-US
;  CURRENT APPLICATION NUMBER: US/10/109,498
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 60/281,261
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: PA 2001 00477
;  PRIOR FILING DATE: 2001-03-22
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(406)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-10-109-498-1

  Query Match             48.7%;  Score 96;  DB 9;  Length 406;
  Best Local Similarity   70.7%;  Pred. No. 5e-08;
  Matches   29;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 ANSFLXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFW 41
          ||:|||||| ||||||| || | ||||:|||:|   |  ||
Db      1 ANAFLXXLRPGSLXRXCKXXQCSFXXARXIFKDAXRTKLFW 41


RESULT 11
US-09-759-130B-313
; Sequence 313, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
```

```
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:  PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:  USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 313
;   LENGTH: 96
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-130B-313

  Query Match             42.9%;  Score 84.5;  DB 9;  Length 96;
  Best Local Similarity   38.6%;  Pred. No. 9e-07;
  Matches   17;  Conservative    8;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2 NSF-LXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           | | |     |:| | |   :|::  |: || | | |:||| ::
Db      36 NRFDLELFTPGNLERECNEELCNYEEAREIFVDEDKTIAFWQEY 79


RESULT 12
US-10-189-123-43
; Sequence 43, Application US/10189123
; Publication No. US20030082586A1
; GENERAL INFORMATION:
;  APPLICANT: KIRST, Susan J.
;  APPLICANT:  HOLTZMAN, Douglas A.
;  APPLICANT:  FRASER, Christopher C.
;  APPLICANT:  SHARP, John D.
;  APPLICANT:  BARNES, Thomas S.
;  TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  FILE REFERENCE: 10147-11U3
;  CURRENT APPLICATION NUMBER: US/10/189,123
;  CURRENT FILING DATE:  2002-07-02
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-06-16
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 100
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
;   LENGTH: 96
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-189-123-43

  Query Match             42.9%;  Score 84.5;  DB 9;  Length 96;
  Best Local Similarity   38.6%;  Pred. No. 9e-07;
  Matches   17;  Conservative    8;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2 NSF-LXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           | | |     |:| | |   :|::  |: || | | |:||| ::
Db      36 NRFDLELFTPGNLERECNEELCNYEEAREIFVDEDKTIAFWQEY 79


RESULT 13
US-09-759-130B-312
; Sequence 312, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:  PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:  USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 312
;   LENGTH: 209
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-130B-312

  Query Match             42.9%;  Score 84.5;  DB 9;  Length 209;
  Best Local Similarity   38.6%;  Pred. No. 2.1e-06;
  Matches   17;  Conservative    8;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2 NSF-LXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           | | |     |:| | |   :|::  |: || | | |:||| ::
Db      36 NRFDLELFTPGNLERECNEELCNYEEAREIFVDEDKTIAFWQEY 79


RESULT 14
US-10-189-123-42
; Sequence 42, Application US/10189123
```

```
; Publication No. US20030082586A1
; GENERAL INFORMATION:
;  APPLICANT: KIRST, Susan J.
;  APPLICANT:  HOLTZMAN, Douglas A.
;  APPLICANT:  FRASER, Christopher C.
;  APPLICANT:  SHARP, John D.
;  APPLICANT:  BARNES, Thomas S.
;  TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  FILE REFERENCE: 10147-11U3
;  CURRENT APPLICATION NUMBER: US/10/189,123
;  CURRENT FILING DATE:  2002-07-02
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-06-16
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 100
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
;   LENGTH: 209
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-189-123-42

  Query Match             42.9%;  Score 84.5;  DB 9;  Length 209;
  Best Local Similarity   38.6%;  Pred. No. 2.1e-06;
  Matches   17;  Conservative    8;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2 NSF-LXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           | | |     |:| | |   :|::  |: || | | |:||| ::
Db      36 NRFDLELFTPGNLERECNEELCNYEEAREIFVDEDKTIAFWQEY 79


RESULT 15
US-09-759-130B-310
; Sequence 310, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  McCarthy, Sean A
;  APPLICANT:  Fraser, Christopher C
;  APPLICANT:  Sharp, John D
;  APPLICANT:  Barnes, Thomas S
;  APPLICANT:  Kirst, Susan J
;  APPLICANT:  Mackay, Charles R
;  APPLICANT:  Myers, Paul S
;  APPLICANT:  Leiby, Kevin R
;  APPLICANT:  Wrighton, Nicolas
;  APPLICANT:  Goodearl, Andrew
;  APPLICANT:  Holtzman, Douglas A
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;  TITLE OF INVENTION:  PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;  TITLE OF INVENTION:  USES.
;  FILE REFERENCE: MPI00-535OMNIM
;  CURRENT APPLICATION NUMBER: US/09/759,130B
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: US 09/479,249
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/559,497
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 09/578,063
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: US 09/333,159
;  PRIOR FILING DATE: 1999-06-14
;  PRIOR APPLICATION NUMBER: US 09/596,194
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 09/342,364
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: US 09/608,452
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/393,996
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 09/602,871
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 09/420,707
;  PRIOR FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 310
;   LENGTH: 226
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-130B-310

  Query Match             42.9%;  Score 84.5;  DB 9;  Length 226;
  Best Local Similarity   38.6%;  Pred. No. 2.3e-06;
  Matches   17;  Conservative    8;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       2 NSF-LXXLRQGSLXRXCIXXICDFXXAKXIFEDVDDTLAFWSKH 44
           | | |     |:| | |   :|::  |: || | | |:||| ::
Db      53 NRFDLELFTPGNLERECNEELCNYEEAREIFVDEDKTIAFWQEY 96


Search completed: May 15, 2003, 13:29:21
Job time : 20 secs
```